

PCT

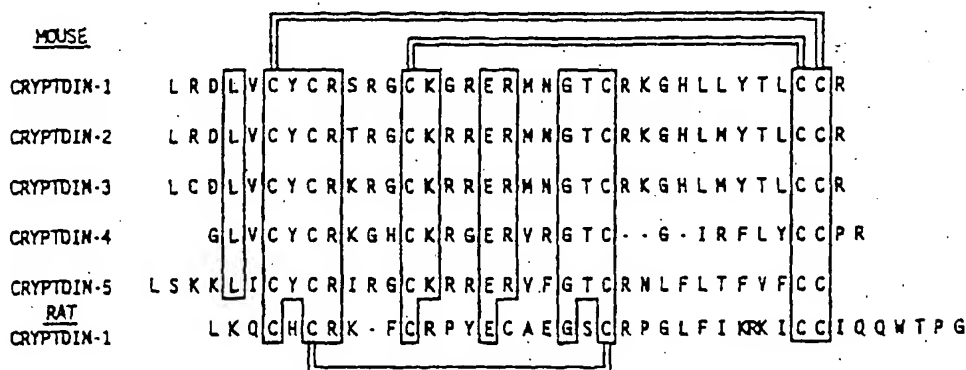
WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: C07H 21/04, C12Q 1/68, G01N 33/53, A61K 37/02, C07K 14/435, 16/18		A1	(11) International Publication Number: WO 96/16075
(21) International Application Number: PCT/US95/13328		(43) International Publication Date: 30 May 1996 (30.05.96)	
(22) International Filing Date: 5 October 1995 (05.10.95)		(81) Designated States: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).	
(30) Priority Data: 08/342,268 18 November 1994 (18.11.94) US		Published With international search report.	
(71) Applicants: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612-3550 (US). THE SHRINER'S HOS- PITAL FOR CRIPPLED CHILDREN [US/US]; P.O. Box 31356, Tampa, FL 33631-3356 (US).			
(72) Inventors: SELSTED, Michael, E.; 16 Young Court, Irvine, CA 92715 (US). OUELLETTE, Andre, J.; 16 Wolcott Road, Lynn, MA 01902 (US).			
(74) Agent: BERLINER, Robert; Robbins, Berliner & Carson, 5th floor, 201 N. Figueroa Street, Los Angeles, CA 90012-2628 (US).			

(54) Title: ANTIBIOTIC CRYPTIDIN PEPTIDES AND METHODS OF THEIR USE



(57) Abstract

The present invention provides substantially purified cryptidin peptides and nucleic acid molecules encoding cryptidin peptides. The invention further provides methods for detecting inflammatory pathologies in a subject and methods for treating an inflammatory pathology in a subject by administering a pharmaceutical composition containing a cryptidin peptide. Representative cryptidin peptides are presented in the figure.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

ANTIBIOTIC CRYPTDIN PEPTIDES AND METHODS OF THEIR USE

This application is a continuation-in-part of United States Serial No. 07/930,649, filed August 14, 1992, which is a continuation-in-part of U.S. Serial No.,
5 07/889,020, filed May 26, 1992, each of which is incorporated herein by reference.

This invention was made with government support under grant numbers AI22931, AI31696, DK08851, DK44632 and DK33506, awarded by National Institutes of Health.
10 The Government has certain rights in the invention.

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

This invention relates generally to
15 antimicrobial peptides and more specifically to cryptdin peptides, nucleic acid molecules encoding cryptdins, and their uses.

BACKGROUND INFORMATION

20 Survival in a world teaming with microorganisms depends on a network of host defense mechanisms. Among these mechanisms are phagocytosis by cells are resident in tissues or that circulate in the blood system and ingest, kill and digest potentially harmful microbes.
25 Although pathogenic microbes may vary considerably, phagocytes are able to destroy the vast majority by sequestering them in intracytoplasmic vacuoles and exposing them to a lethal mixture of organic and inorganic toxins.

30 Perhaps the most remarkable ultrastructural feature of phagocytes are their several thousand cytoplasmic granules, which are membrane-bound organelles typically about 0.3 μ m in diameter. During phagocytosis, some of these granules fuse to phagocytic vesicles thus
35 enabling the contents of the granule to enter the lumen of the vesicle. Early observers surmised correctly that the granules contained factors which were responsible for

intraphagosomal killing in digestion of microbes. These granules contain a mixture of antimicrobial molecules including various peptides such as the so-called defensins.

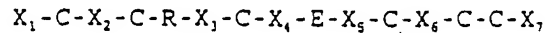
- 5 Defensins are abundant antimicrobial peptide components of vertebrate neutrophil and macrophage granules. Members of the defensin family have been identified previously in human, rabbit, guinea pig and rat phagocytes, primarily those phagocytes termed
10 phagocytic granulocytes. Defensins are cationic peptides that have molecular weights between about 3 and 4 kiloDaltons (kDa) and that exhibit broad-range antimicrobial activities against gram negative and gram positive bacteria, many fungi and some enveloped viruses.
15 The peptides are characterized by eight invariant amino acids, including six invariant cysteine residues that constitute a unique disulfide motif. The three disulfide bonds stabilize a tertiary conformation consisting predominantly of β -sheet. The highly ordered structure
20 and the absence of a helix make defensins unique among known antimicrobial peptides. It appears that defensins exert their antibacterial effect by permeabilizing the cytoplasmic membrane of the target microorganism by a mechanism that may involve the formation of ion channels
25 or transmembrane pores.

- Until recently, defensins had been identified only in circulating or tissue phagocytes of myeloid origin. However, based on the presence of a particular mRNA, it has been surmised that similar peptides might be
30 present in the epithelial cells of the small intestine. Such intestinal peptides may prevent access of microorganisms through the small intestine into the systemic circulation and, therefore, can be useful as a therapeutic or prophylactic agent. Thus, a need exists
35 to identify peptides that have antimicrobial activity within the mucosal epithelium or in the intestinal lumen.

The present invention satisfies this need and provides additional benefits as well.

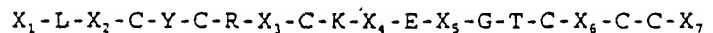
SUMMARY OF THE INVENTION

The present invention provides a substantially purified cryptdin peptide having a consensus amino acid sequence:



wherein X_1 is 3 to 9 amino acids; X_2 is 1 amino acid, preferably Y, H or R; X_3 is 2 or 3 amino acids; X_4 is 3 amino acids; X_5 is 5 amino acids; X_6 is 6 to 10 amino acids; and X_7 is 0 to 9 amino acids.

The invention also provides a substantially purified mouse cryptdin having a consensus amino acid sequence:



wherein X_1 is 3 or 4 amino acids, preferably LRD, LSKK (SEQ ID NO: 8) or LRG;

X_2 is 1 amino acid, preferably V, L or I;

X_3 is 3 amino acids, preferably KGH or *RG,

where * is S, T, K, I or A;

X_4 is 2 amino acids, preferably GR, RR or RG;

X_5 is 3 amino acids, preferably RMN, RVR, RVF HMN or HIN;

X_6 is 6 to 9 amino acids, preferably GIRFLY (SEQ ID NO: 3) or RNLFLTFVF (SEQ ID NO: 4), RRGHLMYTL (SEQ ID NO: 59) or RKGHL*YT* (SEQ ID NO: 5), where * independently is L or M; and

X_7 is 0 to 3 amino acids, preferably R, S or PRR.

For example, the invention provides various mouse, rat or human cryptdins having the sequence:

- 1) LRDLVCYCRSRGCKGRERMNGTCRKGHLLYTLCCR (SEQ ID NO: 9);
- 2) LRDLVCYCRTRGCKRRERMNGTCRKGHLMYTLCCR (SEQ ID NO: 10);
- 3) LRDLVCYCRKRGCKRRERMNGTCRKGHLMYTLCCR (SEQ ID NO: 11);
- 4) GLLCYCRKGHCCKRGERVGTGCGIRFLYCCPR (SEQ ID NO: 12);
- 5) LSKKLICYCRIRGCKRRERVFGTCRNLFITFVFCC (SEQ ID NO: 13);
- 6) LKQCHCRKFRCRPEYKAEGSCRPGLFIKRKICCIQWTPG (SEQ ID NO: 14);
- 7) GLLCYCRKGHCCKRGERVGTGCGIRFLYCCPRR (SEQ ID NO: 15);
- 8) LSKKLICYCRIRGCKRRERVFGTCRNLFITFVFCCS (SEQ ID NO: 16);
- 9) LRDLVCYCRARGCKGRERMNGTCRKGHLLYMLCCR (SEQ ID NO: 17);
- 10) LKQCHCRKFRCRPEYKAEGSCRPGLFIKRKICCIQWTPGRT (SEQ ID NO: 18);
- 11) IGRPVRRRCRCRANGCPKEYATAFCAQGPFKQFKFCCT (SEQ ID NO: 19);
- 12) IRWPWKRCHCRSFCRPPYENATSFCAQGLFKQHKFCCLDTWPPRMK (SEQ ID NO: 20);
- 13) TSGSQARATCYCRTGRCATRESLSGVCEISGRLYRLCCR (SEQ ID NO: 21); and
- 14) AFTCHCRRSCYSTEYSYGTCTVMGINHRFCCL (SEQ ID NO: 22).

20 Cryptdins are typically characterized by being naturally found in the epithelial cells of the small intestine, being cationic, being about 30 to about 45 amino acids in length, having at least three and, preferably, three to nine, amino acids to the N-terminal

25 of the first cysteine residue, exhibiting specific antimicrobial activity against intestinal pathogens and opportunistic pathogens and being relatively non-toxic to cells of the host organism. However, there may be diversity in these structural and functional

30 characteristics. The invention also provides cryptdin analogs, which are devoid of one or more amino acids N-terminal to the first cysteine. In addition, the invention also provides nucleic acid molecules encoding cryptdin peptides. For example, the invention provides

35 genomic DNA sequences and cDNA sequences encoding mouse and rat cryptdins.

The invention further provides a method for detecting an inflammatory pathology in a subject by determining the amount of cryptdin in a biological sample from the subject and comparing that amount to the amount present in a normal subject. Such a method can be used to determine the presence of an inflammatory pathology such as inflammatory bowel disease, pancreatitis, malignancy, infection or ileitis.

The invention also provides a method for treating an inflammatory pathology in a subject by administering a cryptdin to the subject. Such treatment is particularly advantageous in patients who are immunocompromised due, for example, to malnutrition, radiation burns, immunosuppressive infections, autoimmune disease, neonatality, bone marrow transplantation or chemotherapy. A cryptdin can be administered orally, by nasogastric intubation, by transabdominal catheter, intravenously or by aerosol inhalation. When administered orally, it is preferably in a delayed release formulation designed to permit release in the small intestine. The cryptdin can be administered as a composition with a physiologically acceptable medium, and more than one cryptdin can be administered simultaneously or sequentially.

25

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides the structures of mouse cryptdins 1-5 (SEQ ID NOS: 9 to 13, respectively) and rat cryptdin 1 (SEQ ID NO: 14). Amino acid residues are indicated by single letter code. Dashed lines are included in mouse cryptdin 4 (SEQ ID NO: 12) and rat cryptdin 1 (SEQ ID NO: 14) in order to preserve the consensus sequence where these peptides are shorter than other cryptdins. Invariant residues in the enteric cryptdin peptides are boxed. Disulfide bonding motifs are depicted by connecting double lines.

Figures 2.A. to 2.C. show chromatograms representing the purification of enteric cryptdins. Acid extract of jejunum and ileum was chromatographed in 30% acetic acid on a P-60 column. Fractions indicated by the bracket (Figure 2.A.) were pooled and rechromatographed on the P-60 column (Figure 2.B.). Cryptdin containing fractions (bracket, panel B) were pooled and further purified by reversed-phase high performance liquid chromatography (RP-HPLC) on 0.46 x 25 cm Vydac C-18 column. Water-acetonitrile gradient elution (--) using 0.13% (vol/vol) HFBA as modifier was used to purify cryptdins 1-5 (SEQ ID NOS: 9 to 13, respectively). The brackets in Figure 2.C. indicate the peptide contained in each peak, and the portion of each which was subjected to a second round of RP-HPLC.

Figure 3 shows acid-urea PAGE of purified enteric cryptdins. Samples of low molecular weight enteric peptides obtained by P-60 gel filtration (Figure 2, panel B) and purified cryptdins were electrophoresed on a 12.5% acid-urea gel and stained with formalin-containing Coomassie Blue. Lane A: approximately 20 μ g P-60 low molecular weight peptide fractions; lanes B-F: 1 μ g each of cryptdins 1-5 (SEQ ID NOS: 9 to 13), respectively.

Figures 4.A. and 4.B. compare mouse cryptdins 1-5 (SEQ ID NOS: 9 to 13, respectively) and partially purified luminal peptides.

Figure 4.A. Lyophilized luminal lavage of small intestine from 12 mice and 20 μ g protein was fractionated by P-60 gel filtration and electrophoresed on an acid-urea acrylamide gel (lane 2) along side a similarly prepared sample of bowel tissue (lane 1). The positions of cryptdins 1-5 (SEQ ID NOS: 9 to 13, respectively) are indicated.

Figure 4.B. Partially purified luminal peptides (20 μ g; as for Figure 4.A., lane 2) were electrophoresed in a second acid-urea gel (lane 3) along

with an identical sample previously treated with performic acid (lane 4). In lane 4, rapidly migrating, cyst(e)ine-containing peptides are absent due to the increased net negative charge resulting from the conversion of cyst(e)ines to cysteic acid residues.

Figure 5 shows the identification of mouse cryptdins 1-5 (SEQ ID NOS: 9 to 13, respectively) in small intestine epithelium. Acid extracts of intact, whole small intestine (W) or epithelial sheets (E) were lyophilized, dissolved in sample solution and resolved on a 12.5% acid-urea acrylamide gel. Cryptdins 1-5 (SEQ ID NOS: 9 to 13) are identified numerically.

Figures 6.A. to 6.F. show the immunohistochemical localization of cryptdin 1 (SEQ ID NO: 9) in small intestine. Full thickness sections of adult mouse jejunum were incubated with preimmune (Figures 6.A., 6.C. and 6.E.) or anti-cryptdin C rabbit IgG (Figures 6.B., 6.D. and 6.F.) and developed using peroxidase anti-peroxidase secondary antibody magnifications: Figures 6.A. and 6.B., 40X; Figures 6.C. and 6.D., 250X; Figures 6.E. and 6.F., 640X.

Figures 7.A. and 7.B. depict the antimicrobial activity of mouse cryptdin 1 (SEQ ID NO: 9). Samples of purified natural mouse cryptdin 1 (Figure 7.A.) or rabbit NP-1 (Figure 7.B.) were dissolved in 0.01% acetic acid and pipetted into wells produced in a 0.6% agarose/0.3% tryptone plate containing 1×10^6 log phase bacterial cells. After incubation at 37°C for 18 hr, antimicrobial activity was evaluated by measuring the diameters of the clear zones. Closed circles denote wild type *S. typhimurium*; open circles denote the *phoP* mutant.

Figure 8 shows the amino acid sequences for rat cryptdins 1-3 (SEQ ID NOS: 18-20, respectively), human cryptdins 5 and 6 (SEQ ID NOS: 21 and 22; HD-5 and HD-6) and a consensus sequence (Def consensus). Also shown are the amino acids sequences for rat prepro-cryptdins 1-3.

(SEQ ID NOS: 60-62) as deduced from cDNA or genomic DNA sequences as indicated.

Figures 9.A. and 9.B. show the amino acid sequences of mouse cryptdins 1-17 (SEQ ID NOS: 9-11, 15-17 and 23-33, respectively) as determined from the cDNA sequences encoding the specific cryptdin.

Figure 9.A. shows the entire amino acid sequence of the mouse cryptdins. The amino acid sequences of cryptdins 1-6 (SEQ ID NOS: 9-11 and 15-17) were determined by sequencing the purified peptides. The amino acid sequences of cryptdins 7-17 (SEQ ID NOS: 23-33) were deduced from the cDNA sequences (see Figure 10). The amino acids encoded by Exon 1, which encodes the signal peptide and propeptide, and Exon 2, which encodes the mature cryptdin peptide, are indicated. A dot indicates the sequence was not encoded by the cDNA clone. "*" indicates a space, which preserves the homology of the sequences.

Figure 9.B. indicates the degree of relatedness of the mouse cryptdins. Amino acids that are identical to the amino acid shown for cryptdin 1 (SEQ ID NO: 9) are indicated by a dot.

Figure 10 shows the nucleic acid sequences for the cDNA sequences encoding mouse cryptdins 1-17 (SEQ ID NOS: 34-50, respectively). A consensus nucleotide sequence also is shown (SEQ ID NO: 51). A dot indicates the nucleotide is the same as shown for cryptdin 1. The amino acid sequence for cryptdin 1 (SEQ ID NO: 9) is shown above the nucleic acid sequence. Numbers below the nucleotide sequence indicate the nucleotide position relative to the methionine start codon (+1). Numbers above the amino acid sequence indicate the amino acid position. Italics indicate the mature cryptdin peptide sequence. Nucleotides in lower case letters indicate non-coding sequences. "***" indicates a stop codon. "(A)_n" indicates poly-A tail. "*" indicates a space and

"-" indicates the particular nucleotide could not be determined unambiguously.

Figure 11 shows the genomic DNA sequences for mouse cryptdins 1, 2, 3, 5 and 6 (SEQ ID NOS: 53-57, respectively) and the genomic sequence for the apparently inactivated mouse cryptdin i gene (Crypi; SEQ ID NO: 58), in which a stop codon (TGA) is substituted for a cysteine residue. Numbering is as described in the legend to Figure 11. The upper sequence represents a consensus cryptdin gene sequence (SEQ ID NO: 52). "X" indicates positions at which at least two sequences containing nucleotide changes. The TATAAA box is shown in lowercase italics; exons are shown in capital letters; "***" indicates intron DNA; "n" represents approximately 500 base pairs that were not sequenced. "@" indicates the start of the cryptdin peptide coding region at codon 59. Coding sequences are indicated in bold print. Prepro-regions are coded by nucleotides 1-172; cryptdin peptides are coded by nucleotides 173-279. The stop codon is underlined.

GenBank accession numbers for these sequences are 002994 (cryptdin 1, exon 1); 002995 (cryptdin 1, exon 2); 002996 (cryptdin 2, exon 1); 002997 (cryptdin 2, exon 2); 002998 (cryptdin 3, exon 1); 002999 (cryptdin 3, exon 2); 003000 (cryptdin 5, exon 1); 003001 (cryptdin 5, exon 2); 003002 (cryptdin 6, exon 1); 003003 (cryptdin 6, exon 2); 003004 (cryptdin i, exon 1); and 003005 (cryptdin i, exon 2).

Figures 12.A. to 12.C. demonstrate the effectiveness of mouse cryptdins (as indicated) in inhibiting the growth of *E. coli* ML35 cells in an agar diffusion assay.

Figures 13.A. to 13.C. demonstrate the effectiveness of mouse cryptdins (as indicated) in killing *E. coli* cells in suspension.

Figures 14.A. to 14.C. show the cDNA sequences encoding rat cryptdin 1 (Figure 14.A.), rat cryptdin 2

(Figure 14.B.) and rat cryptdin 3 (Figure 14.C.). Nucleotide numbers are indicated.

Figures 15.A. to 15.C. show the genomic DNA sequences encoding rat cryptdin 1 (Figure 15.A.), rat cryptdin 2 (Figure 15.B.) and rat cryptdin 3 (Figure 15.C.). Nucleotide numbers are indicated.

DETAILED DESCRIPTION OF THE INVENTION

The invention provides small peptide molecules, termed cryptdins, which express a broad range of antimicrobial activity, particularly against intestinal pathogens, and for this reason are useful antimicrobial agents. For example, cryptdins have antimicrobial activity against gram negative and gram positive bacteria and against protozoan pathogens (see Example III). Cryptdin peptides and nucleic acid sequences encoding cryptdins were isolated from the small intestine and are active within the epithelial lining of the small intestine and within the lumen of the intestine. Because it is indicative of inflammatory processes, the presence of cryptdins can be utilized in the diagnosis of a wide range of inflammatory conditions.

As used herein, the term "cryptdin" or "enteric defensins" refers to peptides having generally between about 30 and 45 amino acids. Cryptdins are characterized, in part, by a consensus sequence containing six cysteine residues. Illustrative sequences are provided in Figure 1, which shows invariant residues and the disulfide bonding motif. In addition, those residues which are preferably invariant are identified (see, also, Figures 8 and 9).

Cryptdins are further characterized by their cationic charge and their broad range of antimicrobial activity. While related to leukocyte-derived defensins, cryptdins are distinguished from these other molecules by the presence of 3 to 9 amino acids N-terminal to the first cysteine molecule. Cryptdins may have C-terminal

extensions as well. In addition, they exhibit antimicrobial activity against enteric microorganisms, which can become blood-borne pathogens if the intestinal barrier is breached. Since cryptdins are associated with the secretory granules of Paneth cells in the small intestine, they can be secreted from the cells in which they are produced (Sato, Cell Tiss. Res. 251:87-93 (1988); Sato et al., Acta Histochem. 83:185-188 (1988)). Unlike leukocyte-derived defensins, cryptdins are not toxic to mammalian cells.

It should be appreciated that various modifications can be made to the cryptdin amino acid sequence without diminishing the antimicrobial activity of the peptide. It is intended that peptides exhibiting such modifications, including amino acid additions, deletions or substitutions are within the meaning of the term "cryptdin" and, therefore, within the scope of the invention. For example, cryptdin analogs, which are devoid of one or more amino acids N-terminal to the first cysteine residue, are within the present invention. Such cryptdin analogs can be synthesized using well known methods (see Example VI) or can be purified from the intestine where they may occur naturally due, for example, to partial proteolysis of a cryptdin peptide in the intestinal lumen.

Use of the phrase "substantially pure" in the present specification and claims as a modifier of peptide, protein or nucleic acid means that the peptide, protein or nucleic acid so designated has been separated from its *in vivo* cellular environment. As a result of the separation and purification, the substantially pure peptides, proteins and nucleic acids are useful in ways that the non-separated impure peptides, proteins and nucleic acids are not.

The cryptdin peptides of the present invention preferably contain between about 30 and 45 amino acids (see Figures 1, 8 and 9). Cryptdins can be synthesized

by methods well known in the art, such as through the use of automatic peptide synthesizers or by well-known manual methods of peptide synthesis (see Example VI). In addition, they can be purified from natural sources such as small intestinal epithelium of vertebrate, preferably mammalian, origin (see Example I). Such epithelium can be obtained, for example, from rats, mice or humans using means well known to those skilled in the art.

As disclosed herein, various cryptdin peptides were isolated from intestinal epithelium, purified by chromatographic methods and characterized by electrophoresis and amino acid sequencing. Cryptdins were identified by their rapid migration on acid-urea PAGE and by their apparent molecular weight of about 4 kDa (see Examples I and II).

Anti-cryptdin antibodies were made using methods conventional in the art. For example, polyclonal antiserum can be raised in appropriate animals, such as rabbits, mice or rats. Cryptdin peptides, either synthetic or obtained from natural sources, can be used to immunize the animal. As described in Example IV, a cryptdin analog, cryptdin C, which corresponds to residues 4-35 of mouse cryptdin 1 (SEQ ID NO: 9) as shown in Figure 1, was used to immunize rabbits using well known methods. Serum samples were collected until the anti-cryptdin titer was appropriate. Various fractions of the antiserum, such as IgG, can be isolated by means well known in the art. Cryptdin immunogens also can be used to obtain monoclonal antibodies using methods well known in the art (see, for example, Harlow and Lane, Antibodies: A Laboratory Manual (Cold Spring Harbor Laboratory Press 1988), which is incorporated herein by reference).

The antimicrobial activity of a cryptdin can be measured against various pathogens. As disclosed in Example III, various microorganisms were grown to an appropriate concentration, mixed with an appropriate

medium such as an agarose-trypticase soy medium and contacted with a cryptdin. Antimicrobial activity was apparent, for example, from the clear zones that surrounded the cryptdins in an agar diffusion assay. The area of the clear zones was concentration dependent (see Figure 12).

Anti-cryptdin antibodies can be used to determine the presence of cryptdin in a biological sample such as a histological sample. For example, sections of small intestine are fixed by means well known to those skilled in the art and incubated with anti-cryptdin antibodies such as an IgG fraction of antiserum. If desired, the anti-cryptdin antibody can be detectably labelled or an appropriate detectable second antibody can be used to identify the presence of the primary antibody attached to the cryptdin. Means of detection include the use of radioactive protein A or enzyme substrates such as peroxidase (see Harlow and Lane, *supra*, 1988).

Alternative methods of determining the presence of cryptdin in a biological sample obtained, for example, by intestinal lavage or by disrupting cells or tissues can be useful to determine the presence of inflammatory processes. In the presence of inflammatory processes, the concentration of cryptdins is significantly altered from that found in the normal cell. In particular, a deviation from the normal level of cryptdins by one to two standard deviations is indicative of an inflammatory process. Such an inflammatory process can include, for example, inflammatory bowel disease, pancreatitis, malignancy, infection or ileitis.

Because of their broad range of antimicrobial activity and their ability to function within the intestinal epithelium or lumen, cryptdins are potent therapeutic agents for infections of the intestine. In particular, cryptdins are useful where the subject is immunocompromised due, for example, to malignancy, malnutrition, chemotherapy, radiation, immunosuppressive

viruses, autoimmune disease or neonatality. In addition, cryptdins are useful in surgical prophylaxis, for example, by functioning to help sterilize the small bowel. Thus, cryptdins can be useful as medicaments for
5 treating a subject having a pathology characterized, in part, by an inflammatory process.

A cryptdin, either purified from natural sources or synthetic, can be administered to a subject in need of such therapy by various means, including orally,
10 preferably in a slow-release type formulation, which will avoid release within the stomach. Alternatively, cryptdins can be administered through nasogastric intubation, transabdominal catheter, intravenously or aerosol administration. Individual species of cryptdin
15 can be administered alone or in combination. Cryptdins administered in combination can be administered simultaneously or sequentially and can be repeated as necessary.

Prior to the characterization of a mouse
20 intestinal defensin cDNA, expression of defensins was thought to be limited to professional phagocytes, i.e., neutrophils and macrophages. The presence of high levels of cryptdin mRNA in Paneth cells led to the hypothesis that defensins synthesized in intestinal epithelium may
25 contribute to antimicrobial barrier function in the small bowel (Ouellette et al., J. Cell Biol. 108:1687-1695 (1989a), which is incorporated herein by reference). Isolation and characterization of six mouse cryptdin peptides, two rat cryptdin peptides and 2 human cryptdin
30 peptides, and the demonstration of antimicrobial activity of various cryptdin peptides indicates that the cryptdins have an antimicrobial role in the small intestine. The immunohistochemical localization of cryptdin(s) to Paneth cells is consistent with previous *in situ* hybridization
35 analysis and suggests that defensins produced by these cells may contribute to restricting the colonization and invasion of the small bowel by bacteria.

Initial efforts to purify intestinal defensins focused on the isolation of mouse cryptdin 1 (SEQ ID NO: 9), the peptide predicted from the cryptdin cDNA sequence. Since the deduced structure of the peptide is highly cationic, intestinal peptides were solubilized by homogenizing intact mouse jejunum and ileum in 30% formic acid. Acid-urea PAGE of the crude extract revealed several bands with R_f values similar to those of rabbit defensin NP-1 and cryptdin C, a folded synthetic defensin congener corresponding to residues 4 to 35 in cryptdin 1 (SEQ ID NO: 9). Peptides corresponding to these bands were purified approximately 200-fold by sequential gel filtration chromatography on Bio-Gel P-60 (Figures 2.A. and 2.B.). Electrophoresis of P-60 column fractions on acid-urea gels showed that five fractions eluting between two prominent peaks (Figures 2.A. and 2.B., brackets) contained putative cryptdin peptides (Figure 3, lane a). Peptides in these P-60 fractions migrated with an apparent molecular mass of approximately 4 kDa on SDS-PAGE (not shown), consistent with the molecular weight of defensins. Furthermore, treatment of P-60 fraction samples with performic acid reduced the electrophoretic mobility of the five putative mouse cryptdins in acid-urea gels, behavior that is characteristic of defensins and polypeptides that contain multiple cysteine residues.

Defensins in pooled P-60 fractions were purified further using sequential rounds of RP-HPLC utilizing different ion-pair agents. Initial HPLC fractionation utilized water-acetonitrile gradients containing 0.13% heptafluorobutyric acid (HFBA) as the ion-pairing agent, whereby each of the five peptides contained in the pooled P-60 fractions was resolved to near purity in a single run (Figure 2.C.). Complete purification of five peptides, mouse cryptdins 1-5 (SEQ ID NOS: 9 to 13, respectively), was achieved by subsequent RP-HPLC using 0.1% trifluoroacetic acid (TFA) (Figure 3, lanes B-F). Assuming extraction of individual

peptides is equally efficient, both acid-urea gel electrophoresis and RP-HPLC of the P-60 fractions containing putative cryptdins showed that the relative abundance of the peptides is cryptdin 1 > cryptdin 2 > cryptdin 5 > cryptdin 3 > cryptdin 4. The relative amounts of cryptdins 1-5 (SEQ ID NO: 9 to 13, respectively) have been qualitatively reproducible in every preparation of acid-extracted protein from mouse small intestine.

10 Using a modification of the method described above, mouse cryptdin 6, rat cryptdin 2, and human cryptdins 5 and 6 also were isolated (see Examples I and II; see, also, Figures 8 and 9). In addition, longer forms of mouse cryptdins 4 and 5 (compare SEQ ID NOS: 12 and 15; 13 and 16) and rat cryptdin 1 (compare SEQ ID NOS: 14 and 18) were obtained. This result suggests that the initial method of purifying cryptdin peptides resulted in partial degradation of the C-termini of some peptides. Significantly, both forms of the purified
20 cryptdin peptides have antimicrobial activity.

Biochemical characterization of the isolated cryptdins demonstrated that these peptides are defensins. Amino acid analysis of each peptide showed their compositions (cationic peptides of about 30 to 45 amino acid residues, including 6 half-cysteines) are compatible
25 with defensin-like molecules. The complete sequences of mouse cryptdins 1-6 (SEQ ID NOS: 9 to 11 and 15 to 17), rat cryptdins 1 and 2 (SEQ ID NOS: 14, 18 and 19) and human cryptdins 5 and 6 (SEQ ID NOS: 21 and 22) were
30 determined by automated Edman degradation and, in some cases, by amino acid analysis of carboxyl terminal chymotryptic peptides (see Figures 1, 8 and 9). The primary structures of the cryptdins contain the distinctive structural features of human, rabbit, rat and
35 guinea pig neutrophil defensins (Lehrer et al., Cell 64:229-230 (1991a), which is incorporated herein by reference), i.e., the six invariant cysteine residues and

the glycine and glutamic acid in positions that are highly conserved in myeloid defensins.

The cryptdin peptides disclosed herein contain features that are unique and distinct from defensins of myeloid origin. For example, mouse cryptdins 1, 2, 3 and 6 (SEQ ID NOS: 9 to 11 and 17, respectively) are almost identical, differing only at two or three positions (see Figure 9.A.). Analysis of codons from which these amino acid differences could arise shows that the conversion, for example, of Ser¹⁰ to Lys¹⁰ in cryptdin 1 (SEQ ID NO: 9) and cryptdin 3 (SEQ ID NO: 11), respectively, requires two nucleotide substitutions. On the other hand, single nucleotide changes in the codon encoding Thr¹⁰ in cryptdin 2 (SEQ ID NO: 10) could give rise to cryptdins 1, 3 and 6, suggesting that the cryptdin 2 gene may be an intermediate or progenitor of the cryptdin 1, 3 and 6 genes. Similarly, a single nucleotide change in the codon for Thr¹⁰ of cryptdin 2 can account for the deduced amino acid at position 10 in cryptdins 7-17 (see Figure 10, nucleotides 203-205; SEQ ID NOS: 34-50).

By homology with the structures of known myeloid defensins, the cryptdin 1 N-terminus was predicted to begin at Leu⁴ or Val⁵, which is 1-2 residues prior to the first conserved cysteine. However, compared to myeloid defensins, cryptdins have variably extended N-termini that contain from three (mouse cryptdin 4, SEQ ID NO: 12; rat cryptdin 1, SEQ ID NO: 14) to nine (human cryptdin 5, SEQ ID NO: 21) amino acids preceding the first cysteine. In mouse cryptdins 1-3 and 6-17 (SEQ ID NOS: 9 to 11, 17 and 23-33 respectively), the N-peptidyl extensions consist of two charged internal residues flanked by amino acids with hydrophobic sidechains. Since natural variation in defensin amino termini correlates with relative antimicrobial potency *in vitro* (Ganz et al., J. Clin. Invest. 76:1427-1435 (1985), which is incorporated herein by reference), the extended

N-termini of enteric defensins may have evolved for a unique role in the bowel.

Mouse cryptdin 4 (SEQ ID NO: 12), the most cathodal and, apparently, least abundant mouse enteric defensin, was the first defensin found to contain a chain length variation between the fourth and fifth cysteine residues. Unlike the majority of previously known defensins, in which nine amino acids separate the fourth and fifth cysteines (Lehrer et al., *supra*, 1991a), mouse cryptdin 4 (SEQ ID NO: 12) contains only six residues between the same two amino acids (Figure 1). In addition, rat cryptdins 1-3 (SEQ ID NOS: 14 and 18-20) contain ten amino acid residues between the fourth and fifth cysteines. These findings indicate the defensin fold involving this stretch of the peptide chain can accommodate significant variability in the size of the loop, as compared to the invariant loop size defined by crystal and NMR structures, respectively, of human and rabbit neutrophil defensins. Also, rat cryptdins 1-3 (SEQ ID NOS: 14 and 18-20) are the only cryptdins containing three, instead of four, amino acid residues between the second and third cysteine residues.

Since cryptdin mRNA levels increase during postnatal development of mouse small bowel (Ouellette et al., *supra*, 1989a), it was investigated whether accumulation of enteric defensins was regulated similarly. Analysis of intestinal acid extracts from male and female mice showed that mouse cryptdins 1-3 and 5 (SEQ ID NOS: 9 to 11 and 13, respectively) are present in adult mice, regardless of gender. On the other hand, extracts from 9 day-old mice lack the peptides, consistent with postnatal accumulation of cryptdin mRNA.

Mouse cryptdins 1-5 (SEQ ID NOS: 9 to 13) were purified from intestinal epithelial cells. In the presence of EDTA, the intestinal epithelium no longer adheres to the underlying basement membrane and floats free of the lamina propria upon gentle agitation

(Bjerknes and Cheng, Am. J. Anat. 160:51-63 (1981), which is incorporated herein by reference). Preparations of epithelial sheets isolated in this manner were concentrated by low speed centrifugation and extracted with 30% formic acid. Peptides extracted from isolated epithelial sheets comigrate with cryptdins 1-5 (SEQ ID NOS: 9 to 13) when analyzed by acid-urea PAGE (Figure 5), demonstrating their epithelial origin.

Immunoperoxidase staining of full-thickness sections of small intestine with an anti-cryptdin antibody demonstrate cryptdin antigen in Paneth cells, consistent with localization of cryptdin mRNA by *in situ* hybridization (Ouellette et al., *supra*, (1989a)). Incubation of sections of adult mouse jejunum and ileum with a polyclonal anti-cryptdin IgG produced by rabbits immunized with the synthetic congener cryptdin C localized the immunoperoxidase reaction to granulated cells, morphologically defined as Paneth cells, at the base of every crypt (Figure 6). The staining pattern accentuates the granular appearance of the cytoplasm in these cells and the immunoreactivity appears particularly strong over Paneth cell granules. The antibody is specific for mouse cryptdin(s), since it is negative both for rat and human Paneth cells (data not shown). Leukocytes in the lamina propria of the villi also were negative, suggesting that related enteric defensins are not expressed by phagocytes or lymphocytes. Because of the extensive similarity of mouse cryptdins 1-3 (Figure 1; SEQ ID NOS: 9 to 11), the polyclonal antibody produced against cryptdin C probably recognizes the three peptides. Conversely, because mouse cryptdin 4 (SEQ ID NO: 12) and cryptdin 5 (SEQ ID NO: 13) differ markedly from the other mouse cryptdins, the anti-cryptdin C antibody is unlikely to react with cryptdin 4 (SEQ ID NO: 12) and cryptdin 5 (SEQ ID NO: 13), leaving their origin in Paneth cells somewhat unresolved.

Immunohistochemical data suggest cryptdins are secreted into the intestinal lumen. Material in the small intestinal lumen is strongly positive for the antibody but negative for pre-immune sera or IgG (Figures 6.A. and 6.B.). Although the agonist for Paneth cell defensin secretion is unknown, lysozyme, another protein constituent of Paneth cell granules, is secreted into the lumen under cholinergic regulation. Consistent with immunochemical detection of anti-cryptdin C positive material in the intestinal lumen, acid-urea PAGE of saline washes of adult jejunum and ileum contain peptides with mobilities very similar to but distinct from the mobility of cryptdins (Figure 4). Nevertheless, the peptides are not identical to cryptdins 1-5 (SEQ ID NOS: 9 to 13, respectively) by either migration in acid-urea PAGE or by HPLC analysis, suggesting they may correspond to cryptdins that have been processed further. Conceivably, luminal cryptdin or cryptdin-like material could derive from exfoliated Paneth cells in the lumen, but the low rate of Paneth cell turnover suggests this is unlikely. The release of cryptdins or processed variants into the small bowel by Paneth cells contrasts with the apparent lack of defensin secretion by leukocytes, and it is inferred that a secretory pathway may exist for the constitutive delivery of defensins into the intestinal lumen by Paneth cells.

The antibacterial activity of purified mouse cryptdins 1-5 (SEQ ID NOS: 9-13) was tested against wild type and *phoP* mutant *S. typhimurium* using a modified plate diffusion assay (Lehrer et al., J. Immunol. Methods 137:167-173 (1991b), which is incorporated herein by reference). *phoP* is a two-component regulatory locus that is essential to *S. typhimurium* virulence and survival within macrophages (Fields et al., Science 243:1059-1062 (1989); Miller et al., Proc. Natl. Acad. Sci. USA 86:5054-5058 (1989), each of which is incorporated herein by reference). Mutants in the *phoP*

locus are particularly sensitive to rabbit defensins NP-1 and NP-2 when compared to wild type parent strains (Fields et al., *supra*, 1989; Miller et al., Infect. Immun. 58:3706-3710, (1990), which is incorporated herein
5 by reference).

Under assay conditions using a phosphate buffer as described in Example III, the antimicrobial activity of rabbit defensin NP-1 against wild type and the *phoP* mutant organisms was quite similar (Figure 7.B.). On the
10 other hand, at concentrations of mouse cryptdin 1 (SEQ ID NO: 9) that are effective against the attenuated mutant, wild type *S. typhimurium* is completely resistant to the effects of the peptide (Figure 7.A.).

The differential activity of cryptdin 1 (SEQ ID
15 NO: 9) against avirulent *S. typhimurium* suggests that resistance to mucosal defensins may be important for the evolution of virulence in enteric pathogens. However, in experiments using HEPES or PIPES as buffers as described in Example III, concentrations of 100 μ g/ml or 300 μ g/ml
20 cryptdin 1 were as effective as NP-1 in inhibiting the growth of wild type *S. typhimurium*. Furthermore, at these concentrations, cryptdins 4 and 5 were more effective than NP-1 in preventing the growth of mutant and wild type *S. typhimurium* (not shown).

The present invention also provides substantially purified nucleic acid sequences encoding cryptdins. For example, the cDNA sequences for mouse cryptdins 1-17 (SEQ ID NOS: 34-50) are shown in Figure 10 and the cDNA sequences for rat cryptdins 1-3 (SEQ ID NOS:
30 63-65) are shown in Figures 14.A. to 14.C. In addition, the genomic DNA sequences for mouse cryptdins 1, 2, 3, 5 and 6 (SEQ ID NOS: 53-57) and for an apparently inactivated cryptdin gene, cryptdin i (SEQ ID NO: 58) are shown in Figure 11 and the genomic DNA sequences for rat
35 cryptdins 1-3 (SEQ ID NOS: 66-68) are shown in Figures 15.A. to 15.C.

The skilled artisan would recognize that various nucleotide substitutions could be made in the nucleic acid sequences shown in Figures 10, 11, 14 and 15 without altering the amino acid sequence of the encoded cryptdin peptide due to degeneracy of the genetic code. Such nucleotide sequences, which are equivalent to the sequences shown in Figures 10, 11, 14 and 15 are encompassed within the claimed invention.

The invention also provides nucleotide sequences that consist of a portion of a nucleic acid sequence as shown in Figures 10, 11, 14 and 15. Such a nucleotide sequence can be useful, for example, as a probe, which can hybridize under relatively stringent conditions to a nucleic acid molecule encoding a cryptdin peptide. For hybridization, such a nucleotide sequence should be at least about 10 nucleotides in length. One skilled in the art would know that appropriate conditions for hybridization can be determined empirically or can be calculated based, for example, on the G:C content of the nucleotide sequence, the length of the sequence and the number of mismatches, if any, between the probe and the target sequence (see, for example, Sambrook et al., Molecular Cloning: A laboratory manual (Cold Spring Harbor Laboratory Press 1989), which is incorporated herein by reference).

A nucleotide sequence as described above can be detectably labelled by attaching, for example, a radioactive label or biotin, or can be unlabelled. A labelled or unlabelled sequence also can be used as a primer for the polymerase chain reaction (PCR; see, for example, Erlich, PCR Technology: Principles and applications for DNA amplification (Stockton Press 1989), which is incorporated herein by reference). Such a sequence can be useful, for example, to identify a nucleic acid sequence encoding a cryptdin in a cell.

A nucleic acid molecule as shown in Figures 10, 11, 14 and 15 or a nucleotide sequence derived therefrom

also can be useful, for example, for preparing a cryptdin peptide or a portion of a cryptdin peptide using well known methods of recombinant DNA technology. For example, the nucleic acid sequence can be cloned into an expression vector such as a baculovirus vector or a viral vector, which can infect a mammalian cell and express an encoded cryptdin peptide in the cell. Expression from such a vector can be useful for producing large amounts of a cryptdin, which can be used to treat a subject having an inflammatory pathology as described herein, or for producing a cryptdin directly in a subject. Thus, the invention provides vectors containing a nucleic acid molecule as shown in Figures 10, 11, 14 and 15 as well as specific host cells, in which the vector can propagate or can express a cryptdin.

The following examples are intended to illustrate but not limit the invention.

EXAMPLE I

Purification of Enteric Defensins

Outbred Swiss mice, (Crl:CD-1)(ICR)BR, 45 day old males (30-35 g) or timed-pregnant dams, were obtained from Charles River Breeding Laboratories, Inc. (North Wilmington MA). In studies of newborn mice, litters were culled to 8 pups within 12 hr of delivery. Mice were housed under 12 hr cycles of light and darkness and had free access to food and water.

Cryptdins were isolated by a modification of the method described by Selsted et al., J. Cell. Biol. 118:929-936 (1992); Ouellette et al., Infect. Immun. 62:5040-5057 (1994), each of which is incorporated herein by reference. Jejunal and ileal intestinal segments were excised from 60 mice immediately after carbon dioxide euthanasia. The tissue was washed and the lumen was flushed with ice cold water prior to homogenization in 500 ml ice cold 30% acetic acid. The homogenate was clarified by centrifugation, lyophilized, dissolved in

200 ml 30% acetic acid, clarified by filtration through Whatman 541 filter paper and applied to a 10 x 60 cm Bio-Gel P-60 column equilibrated with 30% acetic acid. The elution rate was 100 ml/hr. Fractions containing
5 cryptdins were identified by electrophoresis in acid-urea polyacrylamide gels (Selsted and Harwig, Infect. Immun. 55:2281-2285 (1987), which is incorporated herein by reference).

Cryptdin-containing fractions were pooled and
10 lyophilized, then purification was completed by RP-HPLC. Initial separation of mouse cryptdins 2-5 was achieved by HPLC on a 1 x 25 cm Vydac C-18 column using a gradient of water and acetonitrile containing 0.13% HFBA. Solvents were delivered at 3 ml/min to generate the following
15 acetonitrile gradient: 0-28% (10 min); 28-34% (20 min); and 34-40% (60 min). Cryptdins 1 and 6, which coeluted under these conditions, were resolved by C-18 RP-HPLC using 0.1% TFA as the ion pair and a 16-21% acetonitrile gradient delivered in 35 min at 3 ml/min. To eliminate
20 traces of residual HFBA, preparations of cryptdins 2-5 were subjected to an additional RP-HPLC step using 0.1% TFA. All peptides were lyophilized and quantitated by amino acid analysis prior to antimicrobial testing. Essentially identical methods were used to purify rat and
25 human cryptdin peptides, except that rat cryptdins were isolated from the small intestine of adult Sprague-Dawley rats and human cryptdins were isolated from a surgically resected normal adult human male small intestine.

30

EXAMPLE II

Peptide Characterization

Amino acid analyses were performed on 6 N HCl hydrolysates (150 °C, 2 hr) of unmodified or performic acid-oxidized peptides. Hydrolysates were derivatized
35 with phenylisothiocyanate and the resulting phenylthiocarbamyl amino acids were quantitated as described previously (Selsted and Harwig, *supra*, 1987;

Selsted et al., *supra*, 1992; Ouellette et al., FEBS Lett. 304:146-148 (1992), which is incorporated herein by reference). Peptide samples were reduced with dithiothreitol (DTT) and pyridylethylated with 4-vinyl
5 pyridine for sequencing (Henschen, In Advanced Methods in Protein Microsequence Analysis (Wittmann-Liebold et al., pages 244-255 (1986), which is incorporated herein by reference). Sequence determinations were performed by automated Edman degradation on an ABI model 477 system
10 (Applied Biosystems, Inc.; Foster City CA) with on-line PTH amino acid analysis. In some cases, the C-terminus of a cryptdin peptide was confirmed by amino acid analysis of chymotryptic peptides. Cryptdins 4 and 5 also were analyzed by positive-ion fast atom bombardment
15 mass spectrometry on a VG 7070E-HF instrument (Ouellette et al., *supra*, 1994).

EXAMPLE III

Antimicrobial Assays

20 Antibacterial activity was measured in an agar radial diffusion assay (Lehrer et al., *supra*, 1991b) using wild type *S. typhimurium* (ATCC 10428) or an isogenic phoP mutant of *S. typhimurium* (strain CS015 phoP102::Tn10d-Cam, Miller et al., *supra*, 1989). Cells
25 were grown to log phase in trypticase soy broth at 37 °C, harvested by centrifugation and resuspended to 1×10^7 colony forming units (CFU) per ml in 10 mM sodium phosphate buffer (pH 7.4).

A 100 μ l aliquot of each organism was mixed
30 with 10 ml 1% agarose in 0.03% (w/v) trypticase soy medium, 10 mM sodium phosphate (pH 7.4) at 42 °C. Five μ l samples of peptide solution were pipetted into 3 mm diameter wells formed in the agarose with a sterile punch. After 3 hr at 37 °C, the inoculated agarose plate
35 was overlaid with 1% agarose containing 6% trypticase soy medium. After 12-16 hr, antimicrobial activity was apparent as clear zones surrounding wells loaded with

antibacterial samples; the sizes of the clear zones were concentration-dependent.

Cryptdin antimicrobial activity *in vitro* was substantially enhanced in piperazine-*N,N'*-bis (2-ethane sulfonic acid) (PIPES) or in *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES) as compared to the activity in sodium phosphate. Purified cryptdin peptides were dissolved at 3 to 300 $\mu\text{g/ml}$ in 0.01% acetic acid and activity was examined against *E. coli* ML35 (ATCC). In the radial diffusion assay, 5 μl peptide solution was transferred into wells formed in plates of 1% agarose buffered with 10 mM PIPES (pH 7.4) and containing 1×10^6 log-phase bacteria grown in trypticase soy broth. After 3 hr at 37 °C, the plates were overlaid with 0.8% agarose containing 2X trypticase soy broth and incubated overnight. The antibacterial activities of cryptdin peptides was compared with the activity of rabbit neutrophil defensin NP-1, which was purified from peritoneal exudates as described by Selsted et al. (J. Biol. Chem. 260:4579-4584 (1985), which is incorporated herein by reference). Antibacterial activity was determined by measuring the diameter of clearing around each well and plotted as a function of peptide concentration.

As shown in Figure 12, each cryptdin peptide produced a dose-dependent zone of clearing, which indicates that *E. coli* growth was inhibited. The potencies of the cryptdins varied, with cryptdins 1, 3 and 6 showing similar activity, which was about 3-5x greater than the activity of cryptdin 2. Cryptdin 5 was approximately 5x more active than rabbit NP-1 at concentration above 100 $\mu\text{g/ml}$ (Figure 12.C.) and cryptdin 4 was at least 50x more active than NP-1 when compared at 100 $\mu\text{g/ml}$ and 300 $\mu\text{g/ml}$ (Figure 12.B.). These higher concentrations of cryptdins 4 and 5 also were more effective than the same concentration of NP-1 at inhibiting the growth of *S. aureus* and of wild type

and mutant strains of *S. typhimurium* (not shown). These results demonstrate that various cryptdin peptides can inhibit bacterial growth.

In order to determine whether the effect of the cryptdin peptides against *E. coli* is bacteriostatic or bacteriocidal, bacterial killing was quantitated as a function of time. Bactericidal assays were performed by incubating $1-2 \times 10^6$ log-phase bacteria in 10 mM PIPES containing 10 μ g peptide/ml. After incubation for 15 or 30 min at 37 °C, aliquots were removed, serially diluted and plated on trypticase soy agar. Bactericidal activity was quantitated by counting colonies after overnight incubation at 37 °C.

As shown in Figure 13, cryptdins 1 and 3-6 rapidly killed the *E. coli* cells. In each of these cases, survival was reduced to less than 1% after only 15 min incubation. Cryptdin 2 was the only peptide tested that was not bactericidal under the assay condition. Cryptdins 2 and 3 differ only at amino acid position 10 (threonine and lysine, respectively).

The bactericidal activity of rat cryptdin 1 also was examined. *E. coli* ML35 cells, *S. aureus* 502A cells or mutant or wild type *S. typhimurium* cells were incubated with various concentrations of rat cryptdin 1 or rabbit NP-1. Ten μ g/ml rat cryptdin 1 killed about 90% of the *S. aureus* cells and greater than 99% of the *E. coli* and mutant *S. typhimurium* cells, but was relatively ineffective in killing wild type *S. typhimurium* (not shown). Rat cryptdin 1 was more effective than NP-1 in killing the *E. coli* and mutant *S. typhimurium* cells, whereas NP-1 was more effective in killing *S. aureus*.

The effect of mouse cryptdins 1-3 and 6 at inhibiting the growth of the protozoan, *Giardia lamblia*, which is the most common cause of protozoan disease in the human small intestine, also was examined. Briefly, trophozoites of the C6 clone of *Giardia lamblia* WB (ATCC

30957) were grown to late log phase in TYI-S-33 medium containing bovine bile. Free-swimming trophozoites were discarded and tubes with attached trophozoites were refilled with warm Dulbecco's PBS. Trophozoites were
5 detached by chilling 10 min on ice, then harvested by centrifugation, resuspended at 2×10^7 /ml in 25 mM HEPES (pH 7.5) containing 9% (isotonic) sucrose and incubated for 2 hr at 37 °C with various concentrations of mouse cryptdins 1-3 or 6. Following incubation, trophozoite
10 viability was determined by trypan blue exclusion.

The cryptdin peptides killed *Giardia* trophozoites in a dose-dependent manner (not shown). In particular, 20 µg/ml of cryptdin 2 or cryptdin 3 reduced *Giardia* growth by greater than 2 orders of magnitude (not
15 shown). These results indicate that cryptdins are active against a variety of microorganisms.

EXAMPLE IV

Anti-cryptdin Antibody

20 A polyclonal rabbit antibody was prepared to a synthetic analogue of cryptdin 1. The peptide, termed cryptdin C, corresponding to residues 4-35 in cryptdin 1 (SEQ ID NO: 9; Figure 1) was synthesized by solid phase chemistry using N^o-butoxycarbonyl protection (Kent, Ann.
25 Rev. Biochem. 57:957-989 (1988), which is incorporated herein by reference). Following cleavage/deprotection of synthetic cryptdin C with TFA-trifluoromethanesulfonic acid, the peptide was precipitated in ethyl ether and dried in vacuo. A 100 mg sample was dissolved in 10 ml
30 6.0 M guanidine-HCl, 0.2 M Tris-HCl, pH 8.2, containing 20 mg DTT. The sample was purged with nitrogen, heated to 50 °C for 4 hr and diluted 100-fold with deionized water, then was dialyzed exhaustively, first against 0.1 M sodium phosphate (pH 8.2), 20 mM guanidine-HCl,
35 100 mM NaCl, then against 5% acetic acid. The sample was lyophilized, dissolved in 10 ml 5% acetic acid and subjected to RP-HPLC on a 1 x 25 cm Vydac C-18 column.

The earliest eluting peak, representing about 0.5% of the crude peptide, was determined by amino acid analysis to have the desired composition.

A sample (1.5 mg) of cryptdin C was supplied, without conjugation to carrier, to Berkeley Antibody Company (Berkeley, CA) for immunization of 2 New Zealand White rabbits. Serum samples were collected for 12 weeks, until the anti-cryptdin C titer, determined by ELISA, reached about 1:10,000 for each rabbit. IgG was isolated from antiserum using DEAE Econo-Pac chromatography (Bio-Rad; Richmond CA) as described by the manufacturer.

EXAMPLE V

15

Immunohistochemistry

Paraffin sections of formalin-fixed mouse mid-small bowel were deparaffinized, treated with 1.1% hydrogen peroxide for 40 min, then washed extensively with water followed by PBS. Slides were treated for 20 min at 37 °C with 500 µg/ml trypsin in PBS, washed twice with PBS, and blocked by incubation for 20 min with 5% porcine serum. Slides were incubated for 20 min in rabbit anti-cryptdin IgG (1:10 dilution relative to serum IgG concentration), then washed with blocking serum. Porcine anti-rabbit IgG was used as linking reagent between the primary antibody and rabbit antiperoxidase-
peroxidase conjugate (Dako; Carpinteria CA). Diaminobenzidine was used as peroxidase substrate and parallel incubations were performed using equivalent dilutions of rabbit preimmune IgG as the primary antibody.

EXAMPLE VI

Preparation of Synthetic Cryptdin 1

35

This example provides a method for synthesizing, purifying and characterizing synthetic cryptdin 1.

A. Synthesis

Synthesis was initiated at the 0.13 mmole scale using Wang resin coupled to flourenylmethoxycarbonyl (Fmoc)-arginine using an acid labile linker. Synthesis was carried out in dimethylformamide (DMF) using (relative to resin substitution) a 3-fold excess of Fmoc-amino acids activated in situ with a 3-fold excess of BOP (benzotriazol-1-yl-oxy-tris (dimethylamino) phosphonium hexafluorophosphate) and HOBT (hydroxybenzotriazole) and a 6-fold molar excess of N-methylmorpholine (Nmm). Fmoc removal during synthetic cycles was achieved using cycles of 50% and 25% piperidine in DMF. The side-chain protection scheme utilized the following Fmoc-amino acids: OtBut-aspartic acid, Pmc-arginine, tBut-tyrosine, tBut-serine, Trt-cysteine, tBoc-lysine, OtBut-glutamic acid, Trt-asparagine, tBut-threonine and Trt-histidine.

The peptide chain was assembled in a Synostat batch synthesizer using single couplings at all additions except at leucine and valine which were double coupled.

The cycle sequence is as follows:

1. Wash with DMF 4X for 2 min;
2. Deblock: 50% piperidine 1X for 5 min;
3. Deblock: 25% piperidine 1X for 15 min;
4. Wash with DMF 4X for 2 min;
5. Dissolve amino acids + BOP + HOBT in DMF and transfer to reaction vessel;
6. Add Nmm to RV and mix for 60 min; and
7. Wash with DMF 1X for 2 min.

After coupling of the amino terminal residue, the terminal Fmoc group was removed using the following cycle:

1. Wash with DMF 4X for 2 min;
2. Deblock: 50% piperidine 1X for 5 min;
3. Deblock: 25% piperidine 1X for 15 min;
4. Wash with DMF 4X for 2 min;
5. Wash with dichloromethane 1X for 5 min;

6. Wash with isopropanol 4X for 5 min;
 7. Dry under stream of N₂ 1X for 10-20 min;
and
 8. Dry under vacuum 1X for 12 hr.
- 5 The peptide-resin was weighed to determine mass increase. To cleave and deprotect the peptide-resin, it was first reswelled in dichloromethane, then cleaved and deprotected by addition of reagent R (90% trifluoroacetic acid, 5% thioanisole, 3% ethanedithiol, 2% anisole) at a
- 10 ratio of 10 ml/g peptide-resin. Cleavage/deprotection was carried out under nitrogen for 18 hr at RT.

B. Purification

- 15 The cleavage mixture was separated from resin by filtration through a scintered glass funnel, washed with 1-2 ml fresh reagent R and diluted 5-fold with 50% acetic acid. Glacial acetic acid was added to a final acetic acid concentration of 50%. The resulting solution was extracted 3x with 0.33 vol methylene chloride and the
- 20 aqueous phase was lyophilized to dryness, then dissolved in 50% acetic acid and relyophilized. The extraction and lyophilization steps were repeated 3-4 times, then the dry peptide was dissolved in 30% acetic acid at a concentration of 20 mg/ml and passed over an 800 ml
- 25 Sephadex G-10 column equilibrated in 30% acetic acid. Peptide-containing fractions were pooled, lyophilized, dissolved in 5% acetic acid, then diluted ten-fold with water to a final protein concentration of about 1 mg/ml. The solution was adjusted to pH 8.0 with ammonium
- 30 hydroxide and mixed rapidly with a magnetic stirrer at RT in a beaker open to room air. The pH was adjusted periodically to pH 8.0 over a period of 4 days. The solution was then acidified with acetic acid to pH 3.5 and lyophilized.
- 35 C-18 RP-HPLC using 0.1% TFA-water/acetonitrile gradients was used to purify the folded peptide. Fractions were analyzed on acid-urea gels and compared to

natural cryptdin 1. The yield from an initial crude peptide preparation of 500 mg was approximately 30 mg.

C. Characterization

5 Synthetic cryptdin 1 was compared to natural peptide on analytical RP-HPLC, SDS-PAGE and under three different conditions on acid-urea PAGE. For analysis on acid-urea PAGE, peptide was electrophoresed either without modification, after reduction with DTT or after
10 performic acid oxidation. Under all conditions described, native and synthetic cryptdin 1 behaved identically. The amino acid compositions of natural and synthetic cryptdin 1 were indistinguishable.

15

EXAMPLE VII

Cloning of Nucleic Acid Molecules Encoding Cryptdins

Individual crypts were isolated using a modification of the EDTA elution method of Bjerknes and Cheng, *supra*, 1981, as described by Cano-Gauci et al.,
20 Expt. Cell Res. 208:344-349 (1993), which is incorporated herein by reference. Briefly, the central 10 cm of small intestine from an adult C3H/HeJ mouse was everted on a Buchler gradient-making apparatus, then intact crypts were dislodged by vibration in ice cold 30 mM EDTA in
25 calcium-free, magnesium-free PBS. Isolated crypts were disrupted in a sonicating water bath prior to cDNA synthesis.

The crypt library was constructed by mRNA-directed PCR amplification (Cano-Gauci et al., *supra*,
30 1992). Phage were screened at a density of approximately 300 PFU/dish using the partial cDNA clone, asb4/134, as a probe (Ouellette et al., *supra*, 1989a). Positive phage were collected and denatured plasmid cDNA was sequenced by the dideoxynucleotide termination method using
35 Sequenase™ (U.S. Biochemical Corp.; Cleveland OH). Sequencing primers included T3 and T7 promoter primers and Defcr_{p130}, which is a 16-mer that corresponds to

nucleotides 90-105 in cryptdin 1 mRNA (Huttner et al., Genomics 19: 448-453 (1994), which is incorporated herein by reference). Reaction mixtures were separated by electrophoresis in gels consisting of 5% Long Ranger™ (AT Biochem, Inc.; Malvern PA) and DNA sequence data were analyzed (Ouellette et al., *supra*, 1994). Computations for similarity searches of DNA sequences in nonredundant nucleic acid and protein sequence databases were performed at the National Center for Biotechnology Information with the BLAST network service (Ouellette et al., *supra*, 1994).

A cDNA library also was prepared by amplification of cryptdin mRNA (Huttner et al., *supra*, 1994). Total RNA was isolated from the small intestine of a male 129/SVJ mouse using RNazol™ (Biotecx Lab; Houston TX). First strand cDNA synthesis was performed using the cDNA Cycle Kit (Invitrogen; San Diego CA). Amplification of 5' ends was performed using the 5' RACE method (Frohman et al., Proc. Natl. Acad. Sci., USA 85:8998-9002 (1988), which is incorporated herein by reference) with a reverse primer that was specific for a conserved region of the cryptdin 3'-untranslated sequence (UTS).

Blot hybridization of the PCR products using an oligonucleotide probe specific for the cryptdin prepro-coding region detected a single band. DNA from the band was isolated using the GeneClean II™ kit (Bio101; La Jolla CA), subcloned into the Bluescript II plasmid using the pCR-Script SK(+) cloning kit (Stratagene) and transfected into competent XL-1 Blue cells (Stratagene). Colonies containing cryptdin-related sequences were identified by hybridization to a labelled asb4/134 probe. DNA sequence analysis of the positive clones was performed as described above, except that internal primers were utilized as required.

Using these methods, cDNA sequences encoding 17 distinct mouse cryptdin peptides were identified (Figure

10; SEQ ID NOS: 34-50). The various mouse cryptdin cDNA sequences share 93-100% nucleotide sequence identity with cryptdin 1, except cryptdin 5 and cryptdin 4 share 73% and 69% sequence identity, respectively, with cryptdin 1.

5 The amino acid sequences were deduced from the cDNA sequences for the 17 mouse cryptdins (see Figure 9.A.; SEQ ID NOS: 9-11, 15, 16, 17 and 23-33). As shown in Figure 9.A., the cDNA sequences encode prepro-cryptdin peptides consisting of a signal peptide, a propiece and the cryptdin peptide. The prepro-cryptdins, including the mature cryptdin peptide, share significant amino acid sequence identity with cryptdin 1, although cryptdins 4 and 5 are less homologous (Figure 9.B.). Amino acid variability was most striking at position 10 of the mature cryptdin peptide, where either serine, threonine, alanine, isoleucine or lysine can be found. Interestingly, a single nucleotide change in the sequence of cryptdin 2 can account for each of these amino acids. In addition, position 15 can contain arginine or lysine.

10 The amino acid variability among cryptdin peptides can be involved in conferring different antimicrobial properties to the cryptdins.

 Mouse cryptdin genomic clones also were obtained and sequenced (Huttner et al., supra, 1994).

25 Asb4/134 was used as a probe to screen a custom-made 129/SVJ mouse genomic library constructed in lambda DASH II (Stratagene Cloning Systems, Inc.; La Jolla CA). Approximately 1×10^6 phage were screened in duplicate and 25 positive phage were identified. Ten clones were purified and phage DNA was isolated using Qiagen 100 columns (Qiagen, Inc.; Chatsworth CA). Southern blots of Eco RI-digested DNA from individual phage were hybridized to asb4/134 and hybridizing fragments were subcloned into Bluescript II SKTM (Stratagene) or pUC18 (BRL; Gaithersburg MD) for sequencing.

30 Sequencing was performed as described above, except that primers were selected based on the cryptdin

1 cDNA sequence and with the expectation that mouse cryptdin genes would be structurally homologous to the rabbit MCP-1 and MCP-2 defensin genes (see Huttner et al., *supra*, 1994). DNA sequence data were analyzed using
5 the programs of Staden (Biochem. Soc. Trans. 12:1005-1008 (1984) and the University of Wisconsin Genetics Computer Group (Deveréux et al., Nucl. Acids Res. 12:387-395 (1985)). Searches for homology were performed as described above.

10 As shown in Figure 11, screening of the genomic library produced nucleic acid sequences that contained the complete coding sequences for mouse cryptdins 1, 2, 3, 5 and 6 (SEQ ID NOS: 53-57). In addition, a homologous gene, designated cryptdin i (Cryp*i*; SEQ ID
15 NO: 58), which apparently was inactivated due to a point mutation that changed a cysteine codon to an in-frame stop codon, was isolated. Examination of the nucleic acid sequences revealed that the cryptdin genes contain two exons, the first of which codes for the 5'-UTS and
20 the prepro-coding region and the second of which encodes the mature cryptdin peptide and the 3'-UTS (not shown; but see Figure 11.A.). A similar structure has been described for the human cryptdin genes (Jones and Bevins, J. Biol. Chem. 267:23216-23225 (1992)).

25 Similar methods as described above were used to obtain the cDNA sequences encoding rat cryptdins 1-3 (Figures 14.A. to 14.C.; SEQ ID NOS: 63-65, respectively), except that RNA was obtained from the small intestine of Sprague-Dawley rats. In addition,
30 genomic DNA sequences encoding rat cryptdins 1-3 (Figures 15.A. to 15.C.; SEQ ID NOS: 66-68, respectively) were obtained using methods as described above, except that a genomic library containing Sprague-Dawley DNA cloned in EMBL3 was purchased from Clontech (Palo Alto CA).

35 Although the invention has been described with reference to the disclosed embodiments, it should be understood that various modifications can be made without

departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
- (ii) TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods of Their Use
- (iii) NUMBER OF SEQUENCES: 70
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Robbins, Berliner & Carson
 - (B) STREET: 201 N. Figueroa Street, 5th Floor
 - (C) CITY: Los Angeles
 - (D) STATE: California
 - (E) COUNTRY: USA
 - (F) ZIP: 90012-2628
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/930,649
 - (B) FILING DATE: 14-AUG-1992
- (viii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/889,020
 - (B) FILING DATE: 26-MAY-1992
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Berliner, Robert
 - (B) REGISTRATION NUMBER: 20,121
 - (C) REFERENCE/DOCKET NUMBER: 5555-339C1-XPC
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (213) 977-1001
 - (B) TELEFAX: (213) 977-1003

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Leu Ser Lys Lys
1

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

38

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Gly Ile Arg Phe Leu Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Arg Asn Leu Phe Leu Thr Phe Val Phe
1 5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Arg Arg Gly His Leu Met Tyr Thr Leu
1 5

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Peptide
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "Xaa = Amino acid is
independently L or M."

(ix) FEATURE:

(A) NAME/KEY: Peptide
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "Xaa = Amino acid is
independently L or M"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Arg Lys Gly His Leu Xaa Tyr Thr Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Leu Arg Asp Leu Val Cys Tyr Cys Arg Ser Arg Gly Cys Lys Gly Arg
1 5 10 15

39

Glu Arg Met Asn Gly Thr Cys Arg Lys Gly His Leu Leu Tyr Thr Leu
 20 25 30

Cys Cys Arg
 35

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Leu Arg Asp Leu Val Cys Tyr Cys Arg Thr Arg Gly Cys Lys Arg Arg
 1 5 10 15

Glu Arg Met Asn Gly Thr Cys Arg Lys Gly His Leu Met Tyr Thr Leu
 20 25 30

Cys Cys Arg
 35

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Leu Arg Asp Leu Val Cys Tyr Cys Arg Lys Arg Gly Cys Lys Arg Arg
 1 5 10 15

Glu Arg Met Asn Gly Thr Cys Arg Lys Gly His Leu Met Tyr Thr Leu
 20 25 30

Cys Cys Arg
 35

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Gly Leu Leu Cys Tyr Cys Arg Lys Gly His Cys Lys Arg Gly Glu Arg
 1 5 10 15

Val Arg Gly Thr Cys Gly Ile Arg Phe Leu Tyr Cys Cys Pro Arg
 20 25 30

40

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Leu Ser Lys Lys Leu Ile Cys Tyr Cys Arg Ile Arg Gly Cys Lys Arg
1 5 10 15

Arg Glu Arg Val Phe Gly Thr Cys Arg Asn Leu Phe Leu Thr Phe Val
20 25 30

Phe Cys Cys
35

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Leu Lys Gln Cys His Cys Arg Lys Phe Cys Arg Pro Tyr Glu Lys Ala
1 5 10 15

Glu Gly Ser Cys Arg Pro Gly Leu Phe Ile Lys Arg Lys Ile Cys Cys
20 25 30

Ile Gln Gln Trp Thr Pro Gly
35

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly Leu Leu Cys Tyr Cys Arg Lys Gly His Cys Lys Arg Gly Glu Arg
1 5 10 15

Val Arg Gly Thr Cys Gly Ile Arg Phe Leu Tyr Cys Cys Pro Arg Arg
20 25 30

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Leu Ser Lys Lys Leu Ile Cys Tyr Cys Arg Ile Arg Gly Cys Lys Arg
 1 5 10 15

Arg Glu Arg Val Phe Gly Thr Cys Arg Asn Leu Phe Leu Thr Phe Val
 20 25 30

Phe Cys Cys Ser
 35

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Leu Arg Asp Leu Val Cys Tyr Cys Arg Ala Arg Gly Cys Lys Gly Arg
 1 5 10 15

Glu Arg Met Asn Gly Thr Cys Arg Lys Gly His Leu Leu Tyr Met Leu
 20 25 30

Cys Cys Arg
 35

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Leu Lys Gln Cys His Cys Arg Lys Phe Cys Arg Pro Tyr Glu Lys Ala
 1 5 10 15

Glu Gly Ser Cys Arg Pro Gly Leu Phe Ile Lys Arg Lys Ile Cys Cys
 20 25 30

Ile Gln Gln Trp Thr Pro Gly Arg Thr
 35 40

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Ile Gly Arg Pro Val Arg Arg Cys Arg Cys Arg Ala Asn Cys Gly Pro
 1 5 10 15

Lys Glu Tyr Ala Thr Ala Phe Cys Ala Gln Gly Pro Phe Lys Gln Phe
 20 25 30

Lys Phe Cys Cys Thr
35

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Ile Arg Trp Pro Trp Lys Arg Cys His Cys Arg Ser Phe Cys Arg Pro
1 5 10 15

Tyr Glu Asn Ala Thr Ser Phe Cys Ala Gln Gly Leu Phe Lys Gln His
20 25 30

Lys Phe Cys Cys Leu Asp Thr Trp Pro Pro Arg Met Lys
35 40 45

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Thr Ser Gly Ser Gln Ala Arg Ala Thr Cys Tyr Cys Arg Thr Gly Arg
1 5 10 15

Cys Ala Thr Arg Glu Ser Leu Ser Gly Val Cys Glu Ile Ser Gly Arg
20 25 30

Leu Tyr Arg Leu Cys Cys Arg
35

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Ala Phe Thr Cys His Cys Arg Arg Ser Cys Tyr Ser Thr Glu Tyr Ser
1 5 10 15

Tyr Gly Thr Cys Thr Val Met Gly Ile Asn His Arg Phe Cys Cys Leu
20 25 30

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 103 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

43

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Lys Thr Leu Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
 1 5 10 15

Val Gln Ala Asp Pro Ile Gln Glu Ala Glu Glu Glu Thr Lys Thr Glu
 20 25 30

Glu Gln Pro Ala Asp Glu Asp Gln Asp Val Ser Val Ser Phe Glu Gly
 35 40 45

Pro Glu Pro Ser Ala Leu Gln Asn Leu Glu Ile Gly Trp Pro Leu Lys
 50 55 60

Gln Cys His Cys Arg Lys Phe Cys Arg Pro Tyr Glu Lys Ala Glu Gly
 65 70 75 80

Ser Cys Arg Pro Gly Leu Phe Ile Lys Arg Lys Ile Cys Cys Ile Gln
 85 90 95

Gln Trp Thr Pro Gly Arg Thr
 100

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 96 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Lys Thr Leu Val Leu Leu Ser Ala Leu Val Leu Val Ala Tyr Gln
 1 5 10 15

Val Gln Ala Asp Pro Ile Gln Gly Ala Glu Glu Glu Thr Lys Thr Glu
 20 25 30

Glu Gln Pro Ser Asp Glu Asp Gln Asp Val Ser Val Ser Phe Glu Gly
 35 40 45

Pro Glu Ala Ser Ala Leu Gln Asp Phe Glu Ile Gly Arg Pro Val Arg
 50 55 60

Arg Cys Arg Cys Arg Ala Asn Cys Gly Pro Lys Glu Tyr Ala Thr Ala
 65 70 75 80

Phe Cys Ala Gln Gly Pro Phe Lys Gln Phe Lys Arg Phe Cys Cys Thr
 85 90 95

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 103 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```

Met Lys Thr Leu Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
1           5           10           15

Ile Gln Ala Asp Pro Ile Gln Glu Ala Glu Glu Glu Thr Lys Thr Glu
          20           25           30

Glu Gln Pro Ala Asp Glu Asp Gln Asp Val Ser Val Ser Phe Glu Gly
          35           40           45

Pro Glu Pro Ser Ala Leu Gln Asn Leu Glu Ile Arg Trp Pro Trp Lys
          50           55           60

Arg Cys His Cys Arg Ser Phe Cys Arg Pro Tyr Glu Asn Ala Thr Ser
65           70           75           80

Phe Cys Ala Gln Gly Leu Phe Lys Gln His Lys Phe Cys Cys Leu Asp
          85           90           95

Thr Trp Pro Pro Arg Met Lys
          100

```

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 93 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```

Met Lys Thr Leu Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
1           5           10           15

Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu
          20           25           30

Glu Gln Pro Gly Glu Asp Asp Gln Ala Val Ser Val Ser Phe Gly Asp
          35           40           45

Pro Glu Gly Thr Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys
          50           55           60

Tyr Cys Arg Ser Arg Gly Cys Lys Gly Arg Glu Arg Met Asn Gly Thr
65           70           75           80

Cys Arg Lys Gly His Leu Leu Tyr Thr Leu Cys Cys Arg
          85           90

```

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 93 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```

Met Lys Thr Leu Ile Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
 1           5           10           15

Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu
          20           25           30

Lys Gln Pro Gly Glu Glu Asp Gln Ala Val Ser Val Ser Phe Gly Asp
          35           40           45

Pro Glu Gly Ser Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys
          50           55           60

Tyr Cys Arg Thr Arg Gly Cys Lys Arg Arg Glu Arg Met Asn Gly Thr
          65           70           75           80

Cys Arg Lys Gly His Leu Met Tyr Thr Leu Cys Cys Arg
          85           90

```

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 93 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```

Met Lys Thr Leu Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
 1           5           10           15

Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu
          20           25           30

Glu Gln Pro Gly Glu Asp Asp Gln Ala Val Ser Val Ser Phe Gly Asp
          35           40           45

Pro Glu Gly Ser Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys
          50           55           60

Tyr Cys Arg Lys Arg Gly Cys Lys Arg Arg Glu Arg Met Asn Gly Thr
          65           70           75           80

Cys Arg Lys Gly His Leu Met Tyr Thr Leu Cys Cys Arg
          85           90

```

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 92 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Peptide
 (B) LOCATION: 79
 (D) OTHER INFORMATION: /note= "Xaa = Amino acid is
 independently L or M."

(ix) FEATURE:

46

(A) NAME/KEY: Peptide
 (B) LOCATION: 80
 (D) OTHER INFORMATION: /note= "Xaa = Amino acid is independently L or M."

(ix) FEATURE:

(A) NAME/KEY: Peptide
 (B) LOCATION: 82
 (D) OTHER INFORMATION: /note= "Xaa = Amino acid is independently L or M."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Leu Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln Val Gln Ala
 1 5 10 15

Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu Glu Gln Pro
 20 25 30

Gly Glu Glu Asp Gln Ala Val Ser Ile Ser Phe Gly Gly Gln Glu Gly
 35 40 45

Ser Ala Leu His Glu Lys Ser Leu Arg Gly Leu Leu Cys Tyr Cys Arg
 50 55 60

Lys Gly His Cys Lys Arg Gly Glu Arg Val Arg Gly Thr Cys Xaa Xaa
 65 70 75 80

Gly Xaa Ile Arg Phe Leu Tyr Cys Cys Pro Arg Arg
 85 90

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 93 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met Lys Thr Phe Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
 1 5 10 15

Val Gln Ala Asp Pro Ile His Lys Thr Asp Glu Glu Thr Asn Thr Glu
 20 25 30

Glu Gln Pro Gly Glu Glu Asp Gln Ala Val Ser Ile Ser Phe Gly Gly
 35 40 45

Gln Glu Gly Ser Ala Leu His Glu Glu Leu Ser Lys Lys Leu Ile Cys
 50 55 60

Tyr Cys Arg Ile Arg Gly Cys Lys Arg Arg Glu Arg Val Phe Gly Thr
 65 70 75 80

Cys Arg Asn Leu Phe Leu Thr Phe Val Phe Cys Cys Ser
 85 90

47

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 93 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```

Met Lys Thr Leu Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
1           5           10           15

Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu
                20           25           30

Glu Gln Pro Gly Glu Glu Asp Gln Ala Val Ser Val Ser Phe Gly Asp
35           40           45

Pro Glu Gly Thr Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys
50           55           60

Tyr Cys Arg Ala Arg Gly Cys Lys Gly Arg Glu Arg Met Asn Gly Thr
65           70           75           80

Cys Arg Lys Gly His Leu Leu Tyr Met Leu Cys Cys Arg
                85           90

```

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 93 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```

Met Lys Thr Leu Ile Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
1           5           10           15

Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu
                20           25           30

Glu Gln Pro Gly Glu Asp Asp Gln Ala Val Ser Val Ser Phe Gly Asp
35           40           45

Pro Glu Gly Ser Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys
50           55           60

Tyr Cys Arg Thr Arg Gly Cys Lys Arg Arg Glu His Met Asn Gly Thr
65           70           75           80

Cys Arg Lys Gly His Leu Met Tyr Thr Leu Cys Cys Arg
                85           90

```

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 93 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

48

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Lys Thr Leu Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
 1 5 10 15

Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu
 20 25 30

Glu Gln Pro Gly Glu Glu Asp Gln Ala Val Ser Val Ser Phe Gly Asp
 35 40 45

Pro Glu Gly Ser Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys
 50 55 60

Tyr Cys Arg Lys Arg Gly Cys Lys Arg Arg Glu His Met Asn Gly Thr
 65 70 75 80

Cys Arg Lys Gly His Leu Leu Tyr Met Leu Cys Cys Arg
 85 90

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Leu Ala Phe Gln Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu
 1 5 10 15

Thr Lys Thr Glu Glu Gln Pro Gly Glu Asp Asp Gln Ala Val Ser Val
 20 25 30

Ser Phe Gly Asp Pro Glu Gly Ser Ser Leu Gln Glu Glu Ser Leu Arg
 35 40 45

Asp Leu Val Cys Tyr Cys Arg Lys Arg Gly Cys Lys Arg Arg Glu His
 50 55 60

Met Asn Gly Thr Cys Arg Lys Gly His Leu Met Tyr Thr Leu Cys Cys
 65 70 75 80

Arg

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 92 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Lys Thr Leu Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln Val
 1 5 10 15

49

Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu Glu
20 25 30

Gln Pro Gly Glu Asp Asp Gln Ala Val Ser Val Ser Phe Gly Asp Pro
35 40 45

Glu Gly Ser Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys Tyr
50 55 60

Cys Arg Lys Arg Gly Cys Lys Gly Arg Glu Arg Met Asn Gly Thr Cys
65 70 75 80

Arg Lys Gly His Leu Leu Tyr Thr Leu Cys Cys Arg
85 90

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 85 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Ala Leu Val Leu Leu Ala Phe Gln Val Gln Ala Asp Pro Ile Gln Asn
1 5 10 15

Thr Asp Glu Glu Thr Lys Thr Glu Glu Gln Pro Gly Glu Glu Asp Gln
20 25 30

Ala Val Ser Val Ser Phe Gly Asp Pro Glu Gly Thr Ser Leu Gln Glu
35 40 45

Glu Ser Leu Arg Asp Leu Val Cys Tyr Cys Arg Ser Arg Gly Cys Lys
50 55 60

Gly Arg Glu Arg Met Asn Gly Thr Cys Arg Lys Gly His Leu Leu Tyr
65 70 75 80

Met Leu Cys Cys Arg
85

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 93 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Lys Thr Leu Ile Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
1 5 10 15

Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu
20 25 30

50

Glu Gln Pro Gly Glu Glu Asp Gln Ala Val Ser Val Ser Phe Gly Asp
35 40 45

Pro Glu Gly Thr Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys
50 55 60

Tyr Cys Arg Ala Arg Gly Cys Lys Gly Arg Glu Arg Met Asn Gly Thr
65 70 75 80

Cys Arg Lys Gly His Leu Met Tyr Thr Leu Cys Cys Arg
85 90

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Met Lys Thr Leu Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
1 5 10 15

Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu
20 25 30

Glu Gln Pro Gly Glu Glu Asp Gln Ala Val Ser Val Ser Phe Gly Asp
35 40 45

Pro Glu Gly Thr Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys
50 55 60

Tyr Cys Arg Lys Arg Gly Cys Lys Arg Arg Glu His Met Asn Gly Thr
65 70 75 80

Cys Arg Arg Gly His Leu Met Tyr Thr Leu Cys Cys Arg
85 90

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 85 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Ala Leu Val Leu Leu Ala Phe Gln Val Gln Ala Asp Pro Ile Gln Asn
1 5 10 15

Thr Asp Glu Glu Thr Lys Thr Glu Glu Gln Pro Gly Glu Glu Asp Gln
20 25 30

Ala Val Ser Val Ser Phe Gly Asp Pro Glu Gly Ser Ser Leu Gln Glu
35 40 45

Glu Ser Leu Arg Asp Leu Val Cys Tyr Cys Arg Thr Arg Gly Cys Lys
50 55 60

51

Arg Arg Glu Arg Met Asn Gly Thr Cys Arg Lys Gly His Leu Met His
65 70 75 80

Thr Leu Cys Cys Arg
85

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 93 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Met Lys Thr Leu Val Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
1 5 10 15

Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu
20 25 30

Glu Gln Pro Gly Glu Asp Asp Gln Ala Val Ser Val Ser Phe Gly Asp
35 40 45

Pro Glu Gly Ser Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys
50 55 60

Tyr Cys Arg Lys Arg Gly Cys Lys Arg Arg Glu His Ile Asn Gly Thr
65 70 75 80

Cys Arg Lys Gly His Leu Leu Tyr Met Leu Cys Cys Arg
85 90

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 93 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Met Lys Thr Leu Ile Leu Leu Ser Ala Leu Val Leu Leu Ala Phe Gln
1 5 10 15

Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu Glu Thr Lys Thr Glu
20 25 30

Glu Gln Pro Gly Glu Glu Asp Gln Ala Val Ser Val Ser Phe Gly Asp
35 40 45

Pro Glu Gly Thr Ser Leu Gln Glu Glu Ser Leu Arg Asp Leu Val Cys
50 55 60

Tyr Cys Arg Ser Arg Gly Cys Lys Gly Arg Glu Arg Met Asn Gly Thr
65 70 75 80

52

Cys Arg Lys Gly His Leu Met Tyr Thr Leu Cys Cys Arg
85 90

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 82 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Leu Leu Ala Phe Gln Val Gln Ala Asp Pro Ile Gln Asn Thr Asp Glu
1 5 10 15

Glu Thr Lys Thr Glu Glu Gln Pro Gly Glu Glu Asp Gln Ala Val Ser
20 25 30

Val Ser Phe Gly Asp Pro Glu Gly Thr Ser Leu Gln Glu Ser Leu
35 40 45

Arg Asp Leu Val Cys Tyr Cys Arg Lys Arg Gly Cys Lys Arg Arg Glu
50 55 60

His Met Asn Gly Thr Cys Arg Lys Gly His Leu Leu Tyr Thr Leu Cys
65 70 75 80

Cys Arg

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ACACATTGAG CTCCTGCTCA CCAATCCTCC AGGTGACTCC CAGCCATGAA GACACTAGTC 60
CTCCTCTCTG CCCTTGTCTT GCTGGCCTTC CAGGTCCAGG CTGATCCTAT CCAAAACACA 120
GATGAAGAGA CTAAACTGA GGAGCAGCCA GGGGAAGACG ACCAGGCCGT ATCTGTCTCC 180
TTTGAGAGCC CAGAAGGCAC TTCTCTCAA GAGGAATCGT TGAGAGATCT GGTATGCTAT 240
TGTAGATCAA GAGGCTGCAA AGGAAGAGAA CGCATGAATG GGACCTGCAG AAAGGGTCAT 300
TTATTGTACA CGCTCTGCTG TCGCTGAACA TGGAGACCAC AGAGGACAAG ACCAATCATG 360
GTA CTGAGGC CACTGATGCT GGTGCCTGAT GACCACTTCG CAATAAATTG TTCGCAATAT 420
GC 422

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

53

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```

ACACACTGAG CCGCTACTCA CCAATCCTCC AGGTGACTCC CAGCCATGAA GACACTAATC   60
CTCCTCTCTG CCCTCGTCCT GCTGGCCTTC CAGGTCCAGG CTGATCCTAT CCAAAATACA   120
GATGAAGAGA CTAAACTGA GAAGCAGCCA GGGGAAGAGG ACCAGGCCGT ATCTGTCTCC   180
TTTGGAGACC CAGAAGGCTC TTCTCTTCAA GAGGAATCGT TGAGAGATCT GGTATGCTAT   240
TGTAAGACAA GAGGCTGCAA AAGAAGAGAA CGCATGAATG GGACCTGCAG AAAGGGTCAT   300
TTAATGTACA CGTCTGCTG TCGTGAACA TGGAGACCAC AGAGGACAAG ATGACCATGA   360
GTACTGAGGC CACTGATGCT GGTGCCTGAT GACCACTTCG CAATAAATTG CTTGCAATAT   420
GC                                                                    422

```

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```

ACACATTGGG CTCCTGCTCA CCAATCCTCC AGGTGACTCC CAGCCATGAA GACACTAGTC   60
CTCCTCTCTG CCCTCGTCCT GCTGGCCTTC CAGGTCCAGG CTGATCCTAT CCAAAACACA   120
GATGAAGAGA CTAAACTGA GGAGCAGCCA GGGGAAGACG ACCAGGCCGT ATCTGTCTCC   180
TTTGGAGACC CAGAAGGCTC TTCTCTTCAA GAGGAATCGT TGAGAGATCT GGTATGCTAT   240
TGTAAGAAAA GAGGCTGCAA AAGAAGAGAA CGCATGAATG GGACCTGCAG AAAGGGTCAT   300
TTAATGTACA CACTCTGCTG TCGTGAACA TGGAGACCAC AGAGGACAAG ACGAACATGA   360
GTACTGAGGC CACTGATGCT GGTGCCTGAT GACCACTTCG CAATAAATTG TTGCAATAT   420
GC                                                                    422

```

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 365 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```

ACTAGTCCTC CTCTCTGCCC TCGTCCTGCT GGCCTTCCAG GTCCAGGCTG ATCCTATCCA   60
AAATACAGAT GAAGAGACTA AAATGAGGA GCAGCCAGGG GAAGAGGACC AGGCCGTATC   120
TGTCTCCTTT GGAGACCAG AAGGCTCTGC TCTTCATGAA AAATCTTTGA GAGGTTTGT   180
ATGCTATTGT AGAAAAGGAC ACTGCAAAAG AGGAGAACGA GTTCGTGGGA CTGTGGAAT   240
ACGATTTTGT TACTGCTGCC CCCGCCGCTG AACATGCAGA TGACAAAGAT ATGACAACCA   300
TTGTCTCTGA GGCCGCTGAT GCGGGGCTCT GATGACCACT TCTCAAGAAA TGTTTGCAAT   360
ATGCA                                                                    365

```

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 421 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```

ACACATTGGG CTCCTGCTCA CCAATTCTCC AGGTGACCCC CAGCCATGAA GACATTTGTG 60
CTCCTCTCTG CCCTTGCTCT GCTGGCCTTC CAGGTCCAGG CTGATCCTAT CCACAAAACA 120
GATGAAGAGA CTAATACTGA GGAGCAGCCA GGGGAAGAGG ACCAGGCCGT ATCTATCTCC 180
TTTGGAGGCC AAGAAGGGTC TGCTCTTCAT GAGGAATTGT CAAAAAGCT GATATGCTAT 240
TGTAQAATAA GAGGCTGCAA AAGAAGAGAA CGCGTTTTTG GGACCTGCAG AAATCTTTTT 300
TTAACTTTTG TATTCTGCTG CAGCTGAATA TGCAGATGAC AAAGATATGA CAACCATCAG 360
CACTGAGGCC ACTGATGCTG GGGTCTGATG ATCACCTCGC AATAAATTGT TCGCAATATG 420
C 421

```

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 422 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```

ACACACTGAG CTGCTACTCA CCAATCTCTC AGGTGACTCC CAGCCATGAA GACACTAATC 60
CTCCTCTCTG CCCTCGTCTT GCTGGCCTTC CAGGTCCAGG CTGATCCTAT CCAAAATACA 120
GATGAAGAGA CTAAACTGA GGAGCAGCCA GGGGAAGAGG ACCAGGCCGT ATCTGTCTCC 180
TTTGGAGACC CAGAAGGCAC TTCTCTTCAA GAGGAATCAT TGAGATATCT GGTATGCTAT 240
TGTAAGCAA GAGGCTGCAA AGGAAGAGAA CGCATGAATG GGACCTGCAG AAAGGGTCAT 300
TTATTGTACA TGCTCTGCTG TCGCTGAACA TGGAGACCTC AGAGAACAAG ACGACCATGA 360
GTAAGAGGCC CACTGATGCT GGTGCCTGAT GACCACTTCG CAATACATTG TTCGCAATAT 420
GC 422

```

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 420 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```

ACACTGAGCT GCTACTCACC AATCCTCCAG GTGACTCCCA GCCATGAAGA CACTAATCTC 60
CCTCTCTGCC CTCTGCTCTG TGGCCTTCCA GGTCCAGGCT GATCCTATCC AAAACACAGA 120
TGAAGAGACT AAAACTGAGG AGCAGCCAGG GGAAGACGAC CAGGCCGTAT CTGTCTCCTT 180

```

55

TGGAGACCCA GAAGGCTCTT CTCTTCAAGA GGAATCGTTG AGAGATCTGG TATGCTATTG 240
 TAGAACAAGA GGCTGCAAAA GAAGAGAACA CATGAATGGG ACCTGCAGAA AGGGTCATT 300
 AATGTACACG CTCTGCTGTC GCTGAACATG GAGACCTCAG AGAACAAGAC GACCATGAGT 360
 ACTGAGGCCA CTGATGCTGG TGCCTGATGA CCACTTCGCA ATAAATTGTT CGCAATATGC 420

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 342 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GGTGGCCTTC CAGGTCCAGG CTGATCCTAT CCAAAACACA GATGAAGAGA CTAAGACTGA 60
 GGAGCAGCCA GGGGAAGACG ACCAGGCCGT ATCTGTCTCC TTTGGAGACC CAGAAGGCTC 120
 TTCTCTTCAA GAGGAATCGT TGAGAGATCT GGTATGCTAT TGAGAAAAA GAGGCTGCAA 180
 AAGAAGAGAA CACATGAATG GGACCTGCAG AAAGGGTCAT TTAATGTACA CGCTCTGCTG 240
 TCGCTGAACA TGGAGACCAC AGAGGACAAG ACAAGCATGA GTACTGAGGC CACTGATGCT 300
 GGTGCCTGAT GACCACTTCG CAATAAATTG TTCGCAATAT GC 342

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 377 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

ATGAAGACAC TAGTCCTCT CTCTGCCCTC GTCCTGCTGG CCTTCCAGGT CCAGGCTGAT 60
 CCTATCCAAA ACACAGATGA AGAGACTAAA ACTGAGGAGC AGCCAGGGGA AGAGGACCAG 120
 GCGGTATCTG TCTCCTTTGG AGACCCAGAA GGCTCTTCTC TTCAAGAGGA ATCGTTGAGA 180
 GATCTGGTAT GCTATTGTAG AAAAGAGGCC TGCAAAAGAA GAGAACACAT GAATGGGACC 240
 TGCAGAAAGG GTCATTATT GTACATGCTC TGCTGTCGCT GAACATGGAG ACCACAGAGG 300
 ACAAGATGAA CATGAGTACT GAGGCCACTG ATGCTGGTGC CTGATGACCA CTTGCAATA 360
 AATTGTTCCG AATATGC 377

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 375 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GAAGACACTA GTCTCCTCT CTGCCCTCGT CTTGCTGGCC TTCCAGGTCC AGGCTGATCC 60
 TATCCAAAAC ACAGATGAAG AGACTAAAAC TGAGGAGCAG CCAGGGGAAG ACGACCAGGC 120

56

CGTATCTGTC TCCTTTGGAG ACCCAGAAGG CTCTTCTCTT CAAGAGGAAT CGTTGAGAGA 180
 TCTGGTATGC TATTGTAGAA AAAGAGGCTG CAAAGGAAGA GAACGCATGA ATGGAACCTG 240
 CAGAAAGGGT CATTATTGT ACACGCTCTG CTGTCGCTGA ACATGGAGAC CACAGAGGAC 300
 AAGACGAACA TGAGTACTGA GGCCACTGAT GCTGGTGCCT GATGACCACT TCGCAATAAA 360
 TTGTTCCGAA TATGC 375

(2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 352 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

CCCTCGTCTT GCTGGCCTT CAGGTCCAGG CTGATCCTAT CCAAAACACA GATGAAGAGA 60
 CTAAACTGA GGAGCAGCCA GGGGAAGAGG ACCAGGCCGT ATCTGTCTCC TTGAGAGACC 120
 CAGAAGGCAC TTCTCTTCAA GAGGAATCGT TGAGAGATCT GGTATGCTAT TGTAGATCAA 180
 GAGGCTGCAA AGGAAGAGAA CGCATGAATG GAACCTGCAG AAAGGGTCAT TTATTGTACA 240
 TGCTCTGCTG TCGCTGAACA TGGAGACCAC AGAGAACAAG ACGACCATGA GTACTGAGGC 300
 CACTGATGCT GGTGCCTGAT GACCACTTCG CAATACATTG TTCGCAATAT GC 352

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 422 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

ACCATTTGGG CTCCTGCTCA CCAATCCTCC AGGTGACTCC CAGCCATGAA GACACTAGTC 60
 CTCCTCTCTG CCCTCGTCTT GCTGGCCTT CAGGTCCAGG CTGATCCTAT CCAAAACACA 120
 GATGAAGAGA CTAAACTGA GGAGCAGCCA GGGGAAGAGG ACCAGGCCGT ATCTGTCTCC 180
 TTGAGAGACC CAGAAGGCAC TTCTCTTCAA GAGGAATAGT TGAGAGATCT GGTATGCTAT 240
 TGTAGAGCAA GAGGCTGCAA AGGAAGAGAA CGCATGAATG GGACCTGCAG AAAGGGTCAT 300
 TTAATGTACA CGCTCTGCTG TCGCTGAACA TGGAGACCTC AGAGAACAAG ACGACCATGA 360
 GTACTGAGGC CACTGATGCT GGTGCCTGAT GACCACTTCG CAATAAATTG TTCGCAATAT 420
 GC 422

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 388 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

57

GACTCCCAGG CATGAAGACA CTAGTCCTCC TCTCTGCCCT TGTCTGCTG GCCTTCCAGG 60
 TCCAGGCTGA TCCTATCCAA AACACAGATG AAGAGACTAA AACTGAGGAG CAGCCAGGAG 120
 AAGAGGACCA GGGCGTATCT GTCTCCTTTG GAGACCCAGA AGGCACCTCT CTTCAGAGG 180
 AATCGTTGAG AGATCTGGTA TGCTATTGTA GAAAAAGAGG CTGCAAAAAG AGAGAACACA 240
 TGAATGGGAC CTGCAGAAGG GGTCATTAA GTACACACT CTGCTGTCGC TGAACATGGA 300
 GACCACAGAG GACAAGACGA ACATGAGTAC TGAGGCCACT GATGCTGGTG CCTGATGACC 360
 ACCTCGCAAT AAATTGTTCC CAATATGC 388

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 352 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CCCTCGTCTT GCTGGCCTTC CAGGTCCAGG CTGATCCTAT CCAAAACACA GATGAAGAGA 60
 CTAAACTGA GGAGCAGCCA GGGGAAGAGG ACCAGGCCGT ATCTGTCTCC TTGGAGACC 120
 CAGAAGGCTC TTCTCTCAA GAGGAATCGT TGAGAGATCT GGTATGCTAT TGTAACACA 180
 GAGGCTGCAA AAGAAGAGAA CGCATGAATG GGACCTGCAG AAAGGGTCAT TTAATGCACA 240
 CGCTCTGCTG TCGCTGAACA TGGAGACCAC AGAGGACAAG ACGAGCATGA GTACTGAGGC 300
 CACTGATGCT GGTGCCTGAT GACCACTTCG CAATAAATTG TTGCAAAAT GC 352

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 401 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

CCAATCCTCC CAGTGACTCC CAGGCATGAA GACACTAGTC CTCCTCTCTG CCCTTGTCTT 60
 GCTGGCCTTC CAGGTCCAGG CTGATCCTAT CCAAAACACA GATGAAGAGA CTAAACTGA 120
 GGAGCAGCCA GGGGAAGAGG ACCAGGCCGT ATCTGTCTCC TTGGAGACC CAGAAGGCTC 180
 TTCTCTCAA GAGGAATCGT TGAGAGATCT GGTATGCTAT TGTAGAAAA GAGGCTGCAA 240
 AAGAAGAGAA CACATAAATG GGACCTGCAG AAAGGGTCAT TTATTGTACA CTCCTGCTG 300
 TCGCTGAACA TGGAGACCAC AGAGGACAAG ATGACCATGA GTACTGAGGC CACTGATGCT 360
 GGTGCCTGAT GACCACTCGC AATAAATTGT TCGCAATATG C 401

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 391 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```

GGTGACTCCC AGCCATGAAG AACTAATCC TCCTCTCTGC CCTCGTCTG CTGGCCTTCC 60
AGGTCCAGGC TGATCCTATC CAAAACACAG ATGAAGAGAC TAAACTGAG GAGCAGCCAG 120
GAGAAGAGGA CCAGGCCGTA TCTGTCTCCT TTGGAGACCC AGAAGGCACT TCTCTTCAAG 180
AGGAATCGTT GAGAGATCTG GTATGCTATT GTAGATCAAG AGGCTGCAAA GGAAGAGAAC 240
GCATGAATGG GACCTGCAGA AAGGGTCATT TAATGTACAC GCTCTGCTGT CGCTGAACAT 300
GGAGACCTCA GAGAACAAGA CGACCATGAG TACTGAGGCC ACTGATGCTG GTGCCTGATG 360
ACCACTTCGC AATAAATGT TCGCAATATG C 391

```

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 342 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```

GCTGGCCTTC CAGGTCCAGG CTGATCCTAT CAAAATACA GATGAAGAGA CTAAACTGA 60
GGAGAGCCCA GGAGAAGAGG ACCAGGCCGT ATCTGTCTCC TTGGAGACC CAGAAGGCAC 120
TTCTCTTCAA GAGGAATCGT TGAGAGATCT GGTATGCTAT TGAGAAAAA GAGGCTGCAA 180
AAGAAGAGAA CACATGAATG GGACCTGCAG AAAGGGTCAT TTATTGTACA CGCTCTGCTG 240
TCGCTGAACA TGGAGACCAC AGAGGACAAG ATGACCATGA GTACTGAGGC CACTGATGCT 300
GGTCCCTGAT GACCACCTCG CAATAAATG CTTGCAATAT GC 342

```

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 403 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

```

ACACATGGCT CTCTACCAA TCCTCCAGGT GACTCCAGC CATGAAGACA CTAGTCTTCC 60
TCTCTGCTG TCCTGCTGGC CTTCAGGTC CAGGCTGATC CTATCCAAA CACAGATGAA 120
GAGACTAAAA CTGAGGAGCA GCCAGGGGAA GAGACCAGGC TGTGTCTGTC TCTTTTGGAG 180
ACCCAGAAGG CTTTCTCTC AAGAGGAATC GTTGAGAGAT CTGGTATGCT ATTGTAGAAA 240
GAGGCTGCAA AGAAGAGAAC CATGAATGGG ACCTGCAGAA AGGGTCATT ATGTACAGCT 300
CTGCTGTCCG TGAACATGGA GACCCAGAGA CAAGAACATG AGTACTGAGG CCACTGATGC 360
TGGTGCCTGA TGACCACTTC TCAATAAAT GTTCGCAATA TGC 403

```

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 279..286
- (D) OTHER INFORMATION: /note= "N represents DNA that was not sequenced."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```

TATAAATGCA GGCTGGATAT TCACTGTCCA CACATTGAGC TCCTGCTCAC CAATCCTCCA   60
GGTGACTCCC AGCCATGAAG ACACTAGTCC TCCTCTCTGC CCTTGTCTCG CTGGCCTTCC   120
AGGTCCAGGC TGATCCTATC CAAAACACAG ATGAAGAGAC TAAACTGAG GAGCAGCCAG   180
GAGAAGAGGA CCAGGCGGTA TCTGTCTCCT TTGGAGACCC AGAAGGCACT TCTCTTCAAG   240
AGGAATGTGA GTACTGGTGT CCAGAGTGAT GGATGCTTNN NNNNNNTTTT GTATCTCCAG   300
CGTTGAGAGA TCTGGTATGC TATTGTAGAT CAAGAGGCTG CAAAGGAAGA GAACGCATGA   360
ATGGAACCTG CAGAAAGGT CATTATATGT ACACGCTCTG CTGTCGCTGA ACATGGAGA   419

```

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 279..286
- (D) OTHER INFORMATION: /note= "N represents DNA that was not sequenced."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```

TATAAATGCA GACTGGCTCC TCACTTCCA CACATTGGGC TCCTGCTCAC CAATCCTCCC   60
AGTGACTCCC AGCCATGAAG CCACTTGTC TCCTTTCTGC CCTTGTCTCA CTGTCCTTTC   120
AGGTCCAGGC TGATCCTATC CAAAACACAG ATGAAGAGAC TAAACTGAG GAGCAGTCAG   180
GTGAAGAGGA CCAGGCTGTG TCTGTCTCCT TTGGAGACCG AGAAGGCGCT TCTCTTCAAG   240
AAGAATGTGA GTACTGGTGC CCAGTGTGAT GGATGCTTNN NNNNNNTTTT GTGTCTCCAG   300
CGTTGAGAGA TCTGGTATGC TATTGTAGAA CAAGAGGTTG CAAAGAAGA GAACGCATGA   360
ATGGGACCTG CAGAAAGGT CATTATATGT ACACGCTCTG CTGCCGCTGA ACATGGAGA   419

```

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 279..286
- (D) OTHER INFORMATION: /note= "N represents DNA that was not sequenced."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```

TATAAATGCA GGCTGGATAT TCACCTCCA CACATTGGGC TCCTGCTCAC CAATCCTCCA   60

```

60

GGTGACTCCC AGCCATGAAG ACAC TAGTCC TCTCTCTGC CCTCGTCTG CTGGCCTTCC 120
 AGGTCCAGGC TGATCCTATC CAAAACACAG ATGAAGAGAC TAAACTGAG GAGCAGCCAG 180
 GCGAAGACGA CCAGGCTGTG TCTGTCTCTT TTGGAGACCC AGAAGGCTCT TCTCTTCAAG 240
 AGGAATGTGA GTATTGGTGT CCTGTGTGAT GGATGCTTNN NNNNNNTTTT GTGTCTCCAG 300
 CGTTGAGAGA TCTGGTATGC TATTGTAGAA AAAGAGGCTG CAAAAGAAGA GAACGCATGA 360
 ATGGGACCTG CAGAAAGGGT CATTTAATGT ACACACTCTG CTGTGGCTGA ACATGGAGA 419

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION: 279..286

(D) OTHER INFORMATION: /note= "N represents DNA that was not sequenced."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

TATAAATGCA AGTTGGCTAC TCACCTCTCCA CACATTGGGC TCCTGCTCAC CAATTCTCCA 60
 GGTGACCCCC AGCCATGAAG ACATTGTGCC TCTCTCTGC CCTGTCTCTG CTGGCCTTCC 120
 AGGTCCAGGC TGATCCTATC CAAAAACAG ATGAAGAGAC TAATACTGAG GAGCAGCCAG 180
 GCGAAGAGGA CCAGGCTGTG TCAGTCTCCT TTGGAGGCCA AGAAGGGTCT GCTCTTCAAG 240
 AAGAATGTGA GTAGTGGTAC GCAGTGTGAT GGATGCTTNN NNNNNNTTTT GTGTCTCCAG 300
 TGTCAAAAAA GCTGATATGC TATTGTAGAA TAAGAGGCTG CAAAAGAAGA GAACGCGTTT 360
 TTGGGACCTG CAGAAATCTT TTTTAACTT TCGTATTCTG CTGTAGCTGA ATATGCAGA 419

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION: 279..286

(D) OTHER INFORMATION: /note= "N represents DNA that was not sequenced."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

TATAAATGCA GGCTGGATAT TCACCTCTCCA CACACTGAGC TGCTACTCAC CAATCCTCCA 60
 GGTGACTCCC AGCCATGAAG ACAC TAATCC TCTCTCTGC CCTCGTCTG CTGGCCTTCC 120
 AGGTCCAGGC TGATCCTATC CAAAATACAG ATGAAGAGAC TAAACTGAG GAGCAGCCAG 180
 GCGAAGAGGA CCAGGCTGTG TCTGTCTCTT TTGGAGACCC AGAAGGCACT TCTCTTCAAG 240
 AGGAATGTGA GTACTGGTGT CCAGTGTGAT GGATGCTTNN NNNNNNTTTT GTGTCTCCAG 300
 CATTGAGAGA TCTGGTATGC TATTGTAGAG CAAGAGGCTG CAAAGGAAGA GAACGCATGA 360

ATGGGACCTG CAGAAAGGGT CATTATTGT ACATGCTCTG CTGTCGCTGA ACATGGAGA 419

(2) INFORMATION FOR SEQ ID NO:63:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 419 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION: 279..286
 (D) OTHER INFORMATION: /note= "N represents DNA that was not sequenced."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

ACTTGAGGGT AACAGCCTCT CCCAATTCCA CACATTGAGC TCCTGCTCAC CAATCCTCCA 60
 GGTGACTCCC AGCCATGAAG ACACATGTC TCCTCTCTGC CCTGCCCCTG CTGGCCCTCC 120
 AAGTCCAGGC TGATCCTATC CAAAACACAG ATGAAGAGAC TAAACTGAG GAGCAGCCAG 180
 GGAAAGAAGA CCAAGCTGTT TCTGTCTCCT TTGGAGACCC AGAAGGCTCT TCTCTTCAAG 240
 AGGAATGTGA GTACTGGTGC CCAGTGTGAT GGATGCTTNN NNNNNNTTTT GTGTCTCCAG 300
 CGTTGAGAGA TCTGATATGA TATTGTAGAA CAAGAGGCTG CAAAAGAAGA GAACGCCTGA 360
 ATGGGACCTG AAGAAAGGGT CATTATTGT ACATGCTCTG CTGCTGCTGA ACATGGAGA 419

(2) INFORMATION FOR SEQ ID NO:64:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 411 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

TATAAATGCA RRCTGGMTHY TCACNTTCCA CACATTGRGC TCCTGCTCAC CAATCCTCCA 60
 GGTGACTCCC AGCCATGAAG ACACATGTC TCCTCTCTGC CCTGTCTCTG CTGGCCCTCC 120
 AGGTCCAGGC TGATCCTATC CAAAACACAG ATGAAGAGAC TAAACTGAG GAGCAGCCAG 180
 GGGAAGARGA CCAGGCTGTD TCTGTCTCTT TTGGAGACCV AGAAGGCCCT TCTCTTCAAG 240
 ARGAAATGTGA GTABTGGTGY CCAGTGTGAT GGATGCTTTT TTGTGTCTCC AGCGTTGAGA 300
 GATCTGRTAT GCTATTGTAG ADHAAGAGGC TGCAAARGAA GAGAACGCVT GAATGGGACC 360
 TGCAGAAAGG GTCATTTAAT GTACANNCTC TGCTGYRGCT GAACATGGAG A 411

(2) INFORMATION FOR SEQ ID NO:65:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 445 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

ACACTGGTCT CCAGCTCACC AATCCTCCAG GTGACTTCCA GCCATGAAGA CTCTGTCTCT 60

62

CCTCTCTGCC CTGTCTGCG TGGCATTCCA GGTCCAGGCT GATCCCATTC AAGAGGCAGA 120
 AGAAGAGACT AAAACTGAGG AGCAGCCAGC AGATGAGGAC CAGGATGTGT CTGTCTCCTT 180
 TGAAGGCCCA GAACCTCTG CTCTTCAAAA TTTAGAGATA GGATGGCCAT TAAAGCAGTG 240
 CCATTGCCGA AAGTTCTGCA GACCTTATGA AAAGGCCGAG GGGTCCTGTC GTCCAGGTCT 300
 ATTTATAAAA CGCAAAATCT GCTGCATACA ACAATGGACA CCAGGGAGGA CATAACCAGG 360
 TGAAGTGGGA CCTCACAATC TGTCATTCTT GGGCTTCAAC TCGACTGCTT TTCCTTCTCC 420
 AATAAACCCC TTGCAGACAA AAAAA 445

(2) INFORMATION FOR SEQ ID NO:66:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 445 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ACACTGGTCT CCAGCTCACC AATCCTCCAG GTGACTTCCA GCCATGAAGA CTCTTGTCTT 60
 CCTCTCTGCC CTGTCTCTGG TGGCCTACCA GGTCCAGGCT GATCCCATTC AAGGGGCAGA 120
 AGAAGAGACT AAAACTGAAG AGCAACCATC AGATGAGGAC CAGGATGTGT CTGTCTCCTT 180
 TGAAGGCCCA GAAGCCTCTG CTCTTCAAGA TTTTGAGATA GGAAGGCCAG TGAGGAGGTG 240
 CCGTTGCAGA GCAAACTGCG GACCTAAAGA ATATGCCACT GCGTTCTGTG CTCAGGTCTC 300
 ATTTAAACAG TTCAAATCTT GCTGCACATG AACATGGATC CCAAGTCTGA GATAACCAGG 360
 TGCTCTGGGA CCTCACAATC TGTCATTATT GTGCTTGACC TCAACTGCTT TTCCTTCTCC 420
 AATAAACTCC TGGCAGACAA AAAAA 445

(2) INFORMATION FOR SEQ ID NO:67:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 445 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

ACACTGGTCT CCAGCTCACC AATCCTCCAG GTGACTTCCA GCCATGAAGA CTCTTGTCTT 60
 CCTCTCTGCC CTGTCTCTGC TGGCATTCCA GATCCAGGCT GATCCCATTC AAGAGGCAGA 120
 AGAAGAGACT AAAACTGAGG AGCAGCCAGC AGATGAGGAC CAGGATGTGT CTGTCTCCTT 180
 TGAAGGCCCA GAACCTCTG CTCTTCAAAA TTTAGAGATC AGATGGCCAT GGAAGAGGTG 240
 CCATTGCAGA AGTTTCTGCA GACCTTATGA AAATGCCACT TCGTTCTGTG CTCAGGTCTT 300
 ATTTAAACAA CACAAATCTT GCTGCCTAGA AACATGGCCC CCAAGGATGA AATAACCAGG 360
 TGCTCTGGGA CCTCACAATC TGTCATCATT GTGCTTGGCC TCAACTTCTT TTCCTTCTCC 420
 AATAAACTCC TTGCAGACAA AAAAA 445

(2). INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2457 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

```

CCTGAGACCA ACTCTGTGAT AATCAGAAAA GTCAATAATG TGTCTGAAAT GTAAGGTGTG   60
CTTCTTGACT GATAGTTCTA AGCCTACAGA GAGATTGATG TGGTCATATC CCATTTAACA   120
ATGATATATA TGTAAATAT ATAAAGATAT ATGTATGTTT AGTATGTATG TTCAATATGT   180
ATGTAAATAA TATTCTTGCT GCTTCACTAG CTTTACACA GAGCTGTAAG TAAAAACATT   240
GTAGCCAATG AATAGTATTT ATTAACATGT AAATAGGAGC TGGCACCTGT GACAGTGGGA   300
CTCCATACAC TGACTGTAAA CAACAGGATG CTCTGGACCT TTTGCTGTGT GTGTGGTGAG   360
AGACATGGGA TAAACACAGA CTGAAGAGTG TTCCTGAATG ACATGGCGGC ACTTCTCGAG   420
ACCGGGTAGC AGCTTCTGAG CCTCTCTACA TTGTGGATGT CCTTCTCTGT AGGTCAGGTC   480
TCATTGTCTA AAAGTAAAG CATTGCAGCA TCTCAGACCT GGGAAACACC CCATGGCTTG   540
AGGGTCTCGA GCATGAAGAG CCACCTGGAG CTCACCTCTG GCAGATGTGT TCCATGACTT   600
TGGCTTCTTC AGAACAACCC ACTACAGCTT CACTCTGACA AATCCTAGAA ACTTGAATC   660
AATTCAC TAG AGGGCACCAT AAAGCCATCA TACCTTATAA TGGCCCCAAA GGAGGTGATT   720
CACAAAGTTT GCCTTGATGA GCACAATTGC TAATACACAA AAAGTTGCAA AAAAAAATTG   780
AGTGTCCAGT CCACCTGGTC AAGGACTGGT CCGGATCCA CAGTTTCTGA GAATAGCAGG   840
CTCTAACTTG AAAACACAAA AATTGTTTGT TCTATGAGCT CATTAAATTA GGCAGTGTTC   900
AGCTATTTTC TTTCTGACC ACTGAGAGGT AAATACTCAA GCAGATGGGA AACAGGGGAG   960
GACAGTAAAG CCTGTTTATC ATTATCAGTG GGAGTGTGCA TGAGGGGAGG GGTGTCAGTG  1020
AACACACAGA GCATCAGGAA GGAAGCCTTG AGGACAGAGG AACATCAAG GGATCCTGAG  1080
GACAACAGCT GGGAGCAGTT GCCATCAATG AGTGCCCTCT CTAAGTATGG GGCATGTTCT  1140
TTGCCCTATA AATGCAGGCT GGCTTCTCTC TCCACACACT GGTCTCCAGC TCACCAATCC  1200
TCCAGGTGAC TTCCAGCCAT GAAGACTCTT GTCTCTCTCT CTGCCCTTGT CCTGCTGGCA  1260
TTCCAGGTCC AGGCTGATCC CATTCAAGAG GCAGAAGAAG AGACTAAAC TGAGGAGCAG  1320
CCAGCAGATG AGGACCAGGA TGTGTCTGTC TCCTTTGAAG GCCCAGAAAC CTCTGCTCTT  1380
CAAAATTTAG GTGCGTGCTT GTGCACAGAA TGATGGAGGC TTGGAGTCTC CTGATGGAGG  1440
GTTGTAGATT AGCCCTGGAG TCCTGTCAAG GACAGTCTGG TTCAGGTAGC TGTCTACTGA  1500
TCCTTTTACA ACTTCCCTGT CTTATTCATA GAAATAACAG TGAGAGACAA GCCATTGGGC  1560
TTGACTTTTT CTTTTAAGA TTTCGGTCTA ACAATTTATC TGTGAAAAAC CTTTAAATA  1620
TAAACATAT TGATTAGTTC TTTAAACCTG AGTGATAATT TTCTTACAGG AAGAAATATC  1680
CGTTTTACCC TAAAAATTAG ATTGGTACCC AAATGCCAGT GTATGAAGGT GTTGGGTCAA  1740
GAAAACACAA AAAAAGTGT AGAATATGGT GTAGATGAAA ATTCCTATAT GTGATTAACA  1800
CTTGTTAAAC ATCTTATCTC CATGTGTTTG GGGTTGATCA CTGTCTGGC TGTGATGTCA  1860

```

64

CCCACACAGC AACCTACTC TCTACCATGC ACAGGACATC TCATGGGGT AGTTCAGTGT 1920
 TACACACTAC TGGGCTCCTT ACTTCATGCC TGATGCTTTC TTGTTTCCTC AGAGATAGGA 1980
 TGGCCATTAA AGCAGTGCCA TTGCCGAAAG TTCTGCAGAC CTTATGAAAA GGCCGAGGGG 2040
 TCCTGTCTGC CAGGTCTATT TATAAACGC AAAATCTGCT GCATACAACA ATGGACACCA 2100
 GGGAGGACAT AACCACTGA ACTGGGACCT CACAATCTGT CATTCTTGGG CTTCAACTCG 2160
 ACTGCTTTTC CTTCTCCAAT AAACCCCTTG CAGACAAATA ACCTGTTTAT GTTTTTTTGA 2220
 TGCTTTCTAT GTGGCGTAGA CAGGACTCTC CTGAGCCATG TAGCAAAATC TTCAGTGAAT 2280
 CCTTTGTAAA AGAAGTCTTG GTCACATTC AGCAGTCATA TCAAGGATGA GCAGGAGGTT 2340
 AGATCCAAAG AGACAAGATG GTCTGGGCA GCTGCTTCTG TGTCTATCAA GTCTTCTGTC 2400
 CTTTAGATTA GAGTCACCT CAAAAATTAG TTCCAGATT TCATGTTCTA TTTTTTC 2457

(2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2408 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TATTACGAAT TCGAGCTCGG TACCGGTATA TGAAGAGCGA CCACTGCCAG GACGAAAGTG 60
 CAATGCGGCA TACCTCAGTG CCGTGGAGTG CAGGTATACA GATTAATCCG GCAGCGTCCG 120
 TCGTTGTTGA TATTGCTTAT GAAGGCTCCG GCAGTGGCGA CTGGCGTACT GACGGATTCA 180
 TCGTTGGGGT CGGTATATAA TTCTGATTAG CCAGGTAACA CAGTGTATG ACAGCCCGCC 240
 GGAACCGGTG GCGTTTTTGG TGGGGTGAAT ATGGCAGTAA AGATTTCAGG AGTCTGAAA 300
 GACGGCACAG GAAACCGGT ACAGAACTGC ACCATTGAGC TGAAGGCCAG ACGTAACAGC 360
 ACCACGGTGG TGGTAACAC GGTGGGCTCA GAGAATCCGG ATGAAGCCTG CTTTTTTATA 420
 CTAAGTTGGC ATTATAAAAA AGCATTGCTT ATCAATTTGT TGCAACGAAC AGGTCACTAT 480
 CAGTCAAAAT AAAATCATT TTTGATTTCA ATTTTGTCCC ACTCCCTGCC TCTGTCATCA 540
 CGATACTGTG ATGCCATGGT GTCCGACTTA TGCCCGAGAA GATGTTGAGC AAACCTATCG 600
 CTTATCTGCT TCTCATAGAG TCTTGCAGAC AAACCTGCCA ACTCGTGAAA GGTAGGCGGA 660
 TCTGGGTCCA CTCTAGGCCT CACTGGCCTA ATACGACTCA CTATAGGGAG CTCGAGGATC 720
 ATTGCTAATA CCATGAACT TGACCACCTG GTCAAGGACT GGTCCAGGGT CCACAGTTTC 780
 TGAGAAGAGC AGGCTCCAAC TTCTAACCA AAAAATATT TTTTCCATGC GTCCTTAAA 840
 TTAGGCAGCG CCCAGCTATT TTCTTTCTG ACCACTGAGA GGTAATACT CAAGCAGATG 900
 GGAAACAGGG GAAGATAGCA AGGCCTCTTC ATCATTATCA CTGGGTGTGT GCGTGAGGGG 960
 AGGGGTGTCA TTGCATACAC AGGGCAACAT CAGGATGGAA GCCTTGAGGA CAGAGGAACA 1020
 TCAAAGGGAT CCTGAGGACA ACAGCTGGGA GCAGTTGCCA TCAGTGAGTG CCTTCTCTAA 1080
 GTGTGGGGCC TTTCTCTGCC ACATAAATGC AGGCTGCCTC CTCTCTCCAC AACTGGTCT 1140
 CCAGCTCACC AATCCTCCAG GTGACTTCCA GCCATGAAGA CTCTTGCTCT CCTCTCTGCC 1200
 CTTGTCTCTG TGGCCTACCA GGTCCAGGCT GATCCCATTC AAGGGGCAGA AGAAGAGACT 1260

65

AAAACTGAAG AGCAACCATC AGATGAGGAC CAGGATGTGT CTGTCTCCTT TGAAGGCCCA 1320
 GAAGCCTCTG CTCTTCAAGA TTTTGGTGAG TGCTTATGCA CAGAAATGATG GAGGCTTGGA 1380
 GTCTCCTGAT GGAGGGTTGT AGATTAGACC TGAATCCTG TCAAGAACTG TCTGGTTGAG 1440
 GTAGCTGTCT CTTGGTCCCT TTACATTCCT TGTCTTCTTC ATAGAAGTAA CGGAGAGAGA 1500
 TTAACCATTG GGCTTGACTT TTTTCCTTTT AAAATTTTGT ATCTAACAAAT TTATCTGTGG 1560
 AAAACCTTTA AAATATAAAA CATATTGATT AGTTCTTTTA GACCTGATTG ATAATTTTGT 1620
 TATAAGAAGA AATATTCGTT CTACTTTAAA AATTAGATTT GGGACCCAAA TGCCAGTGTA 1680
 TGAAGCTGTT GGGTAAGGAA AAACCAAAAA TGGTGATAGA ATGTTGTGTA GATGACAATT 1740
 CCTTTATGCG ATTAACACTT TTTAAAATGT CTTATCTCCA TGTGTTGGG GTTGATCATG 1800
 GTGCTGACTG TGATGTCACC CACAGAGCAA ACCTACTCTC TACCATGCAC AGGACATCTT 1860
 CATAGGGTAG TTCACTGTCA CACACTGCTG GCCTCGTTAC TTCATGCCTG ATGCTTTCTT 1920
 GTTTCCTCAG AGATAGGAAG GCCAGTGAGG AGGTGCCGTT GCAGAGCAAA CTGGCGACCT 1980
 AAAGAATATG CCACTGGGTT CTGTGCTCAA GGTCCATTTA AACAGTTCAA ATTCTGTGTC 2040
 ACATGAACAT GGATCCCAAG TCTGAGATAA CCACGTGCTC TGGGACCTCA CAATCTGTCA 2100
 TTATTGTGCT TGACCTCAAC TGCTTTTCTT TCTCCAATAA ACTCCTGGCA GACAAATAAT 2160
 CGGTATATGT TTATTGATG CTTTCTATTT GGCTTAGACA GAACTCTCCT GAGCCATGTA 2220
 GCTGAATCTT CAGTGAATCC TTTGTAAAGG TCACATTTC AAGTGCATAT CAAGGATGAG 2280
 CAGGAGGTTA GATACAAAGA GACAAGATGG TCTGCCCCAG CTGCTTCTTT GTCTATCAAG 2340
 TCTGCTTTCC TTTAGATTAG AGTCACCATC AAAAATTATT CCCACATTTT CATGTTCTAT 2400
 ATTTTTTT 2408

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2551 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CCTGAGACCA ACTCTGTGAT AATCAGAAAA GACAATTATG TGTCTTAAAT GTAAGGTTTG 60
 CTTCTTGACT GATAGATCTA ACCCTACAGA GAGATTCAAG TGGCTTTGTC CCATTGAACA 120
 ATAGTATATA TGTTTTATAT ATATATATAT ATATATGTAT ATGTATATAT ATATGTGTGT 180
 GTGTGTGTGT GTGTGTCTGT GTCTGTGTGT CTGTGTGTCT GTGTGTCTGT GTGTCTGTGT 240
 GTGTATGTGT GTGTATGTGT ACATATGTTT AATATGTCTG TAAATAGTA TTCTTGTAGC 300
 TTCCTTACT TTTGCACAGA GCTGTAAATA AGAACATTGT AGCCAATGAA TAGTATTTAT 360
 TAACATGTAA ATAGGAGCTG GCACCTCTGA CAGTGGGACT CCATACAGTG ACTGTAAACA 420
 ACAGGATGCT CTAGACCTTT TGCTGTGTGT GTGGTGAGAG ACATGGGATA AACACAGACT 480
 GAAGTGTATG ACATGGCGCG ACTTCTCGAG ACCGGGTAGC AGCTTCTGAG CCTCTCTACA 540
 TTGTGGATGT CCTTTCCTGT AGGTCAGGTC TCATTGTCTA AAAGTAAAG CATTGCAGCA 600
 TCTCAGACCT GGGAAACACC CCATGGCTTG AGGGTCCCGC AGGTGAAGAG CCACCTGGAG 660

66

CTCACCTCTG GCAGATGTGT TCCATGACTT TGGCTTCTTC AGAACCACCC ACTACAGCTT 720
 CACTCTGACA AATCTTAGAA ACTTGAATC AATTCACCTG AGGGCACAAT AAAGCCATCT 780
 TACTTTCTCT AAAATGGCCC CAAAGGAGGG GATTCACAAA GTTTGCCTTG ATGAGGACCA 840
 TTGCTAATAC CCAAAAACCT GCAAAAAAAA TTGAGTGTC AGTCAACCTG GTCAAGGACT 900
 GGTCTGGAT CCACAGTTTC TGAGAAAAA AGGCTCCAAC TTCAAAACAC AAACCACTCC 960
 TGTCTATGC GCTCATTAAA TTAGGCAGTG TTAAGCTATT TTCCTTCCTG ACCACTGAGA 1020
 GGTAAATACT CAAGCAGATG GGAACAGGG GAGGACAGCA AAGCCTGTTC ATCATTATCA 1080
 GTGGGAGTGT GCGTGAGGG AGGGGTGTCA GTGAACACAC AGAGCATCAG GAAGGAAGCC 1140
 TTGAGGACAG AGGAACATCA AAGGGATCCT GAGGACAACA GCTGGGAGCA GTTGGCATCA 1200
 CTGAGTGCCG TCTCTAAGTG TGGGGCCTTT CTCTGCCACA TAAATGCAGG CTGGCTCCTC 1260
 TCTCCACACA CTGGTCTCCA GCTCACCAAT CCTCCAGGTG ACTTCCAGCC ATGAAGACTC 1320
 TTGCTCTCT CTCTGCCCTT GTCCTGCTGG CATTCAGAT CCAGGCTGAT CCCATTCAAG 1380
 AGGCAGAAGA AGAGACTAAA ACTGAGGAGC AGCCAGCAGA TGAGGACCAG GATGTGTCTG 1440
 TCTCCTTTGA AGGCCAGAA CCCTCTGCTC TTCAAAATTT AGGTGCGTGC TTGTGCACAG 1500
 AATGATGGAG GCTTGGAGTC TCCGTAGCA GGGTTGTAGA TTAGCCCTGG AGTCTGTCA 1560
 AGGACAGTCT GGTTCAGGTA GCTGTCTATT GATCCTTTCA GAACTTCCCT GTCTTATCA 1620
 TAGAAATAAC AGTGAGAGAC AAGCCATTGG GCTTGACTTT TTCCTTTTAA GATTTTGGTC 1680
 TAACAATTTA TCTGTGAAA ACCTTTAAAA TATAAACAT ATTGATTAGT TCTTTTAAAC 1740
 CTGATTGATA ATTTTGTAT AGGAAGAAAT AACTGTTCTA CTTTAAAAAT TAGATTGGT 1800
 ACCTAAATGC CAGTGATTA AGGTGTTGGG TCAGGAAAAC ACAATAATGC TGATAGAATG 1860
 TGGTGTAGAT GACAATTCCT ATATGCGATT AACACTTGT AAATTGTCCT ATCTCCATGT 1920
 GTTTGGGGT GATCATGGTG CTGGCTGTGA TGTCACCCAC ACAGCAAACC TACTTTCTAC 1980
 CATGCACAGG ACATCTTCAT AGGGTAGTTC ACTGTACAC ACTGCTGGCC TCTTACTTC 2040
 ATGCCGTATG CTTTCTCGTT TCTCAGAGA TCAGATGGCC ATGGAAGAGG TGCCATTGCA 2100
 GAAGTTTCTG CAGACCTTAT GAAATGCCA CTTCTTCTG TGCTCAAGGT CTATTTAAAC 2160
 AACACAAAT CTGCTGCTA GAAACATGGC CCCCAGGAT GAAATAACCA CGTGCTCTGG 2220
 GACCTCACA TCTGTATCA TTGTGCTTGG CCTCAACTTC TTTTCTTCT CCAATAAAT 2280
 CCTTGACAC AAATAACCTG TTTATGTTTT TTTGATGCTT TCTATGTGGC TTAGACAGGG 2340
 CTCTCCTGAG CCATGTAGCA GAATCTTCAG TGAATCCTTT GTAAAAGAAG TCTTGGTCAC 2400
 ATTTCAACAG TCATATCAAG GATGAGCAGG AGGTTAGATC CAAAGAGACA AGATGCTCTG 2460
 CTCCAGCTGC TTCTTGACTA TCAAGTCTTC TGTCCTTCAG ATTAGATCA CCCTCAAAAA 2520
 TTAGTCCAC CTTTTCATGT TCTATTTTT T 2551

WE CLAIM:

1. A substantially purified cryptdin peptide of enteric origin having an amino acid sequence as follows:

5 $X_1-C-X_2-C-R-X_3-C-X_4-E-X_5-C-X_6-C-C-X_7$,

wherein X_1 is 3 to 9 amino acids;

X_2 is 1 amino acid;

X_3 is 2 or 3 amino acids;

X_4 is 3 amino acids;

10 X_5 is 5 amino acids;

X_6 is 6 to 10 amino acids; and

X_7 is 0 to 9 amino acids.

2. A substantially purified mouse cryptdin peptide of enteric origin having an amino acid sequence as follows:

$X_1-L-X_2-C-Y-C-R-X_3-C-K-X_4-E-X_5-G-T-C-X_6-C-C-X_7$,

wherein X_1 is 3 or 4 amino acids, preferably LRD, LSKK (SEQ ID NO: 8) or LRG;

20 X_2 is 1 amino acid, preferably V, L or I;

X_3 is 3 amino acids, preferably KGH or *RG,

where * is S, T, K, I or A;

X_4 is 2 amino acids, preferably GR, RR or RG;

X_5 is 3 amino acids, preferably RMN, RVR, RVF

25 HMN or HIN;

X_6 is 6 to 9 amino acids, preferably GIRFLY (SEQ ID NO: 3), RRGHLMYTL (SEQ ID NO: 59) or RNLFLTFFV (SEQ ID NO: 4) or RKGHL*YT* (SEQ ID NO: 5), where * independently is L or M; and

30 X_7 is 0 to 3 amino acids, preferably R, S or PRR.

3. The substantially purified cryptdin peptide of claim 2, wherein X_1 is selected from the groups consisting of LRD, G and LSKK (SEQ ID NO: 8).

4. The substantially purified cryptdin of claim 2, wherein X_2 is selected from the group consisting of V, L and I.

5. The substantially purified cryptdin of claim 2, wherein X_3 is selected from the group consisting of KGH and *RG, wherein * is selected from the group consisting of S, T, K, I and A.

6. The substantially purified cryptdin of claim 2, wherein X_4 is selected from the group consisting of GR, RR and RG.

7. The substantially purified cryptdin of claim 2, wherein X_5 is selected from the group consisting of RMN, RVR, RVF, HMN and HIN.

8. The substantially purified cryptdin of claim 2, wherein X_6 is selected from the group consisting of GIRFLY (SEQ ID NO: 3), RNLFLTFVF (SEQ ID NO: 4), RRGHLMYTL (SEQ ID NO: 59) and RKGHL*YT* (SEQ ID NO: 5), wherein * indicates L or M independently.

9. The substantially purified cryptdin of claim 2, wherein X_7 is selected from the group consisting of R, S and PRR.

10. The substantially purified cryptdin of claim 2, wherein the amino acid sequences X_1 , L, X_2 are absent.

11. A substantially purified cryptdin peptide of enteric origin having an amino acid sequence selected from the group consisting of:

- 5 GLLCYCRKGHCCKRGERVRGTCGIRFLYCCPRR (SEQ ID NO: 15);
 LSKKLICYCRIRGCKRRERVFGTCRNLFLTFVFCCS (SEQ ID NO: 16);
 LRDLVCYCRARGCKGRERMNGTCRKGHLLYMLCCR (SEQ ID NO: 17);
 LKQCHCRKFCRPYEKAEGSCRPGFLIKRKICCIQQWTPGRT (SEQ ID
 NO: 18);
10 IGRPVRRRCRANCGPKEYATAFCAQGPFKQFKFCCT (SEQ ID NO: 19);
 IRWPWKRCHCRSFCRPYENATSFCAQGLFKQHKFCCLDTWPPRMK (SEQ ID
 NO: 20);
 TSGSQARATCYCRTGRCATRESLSGVCEISGRLYRLCCR (SEQ ID
 NO: 21); and
15 AFTCHCRRSCYSTEYSYGTCTVMGINHRFCCL (SEQ ID NO: 22).

12. A pharmaceutical composition, comprising a cryptdin peptide and a physiologically acceptable carrier.

20

13. A method for detecting an inflammatory pathology in a subject, comprising the steps of:

- a. determining the amount of a cryptdin
 in a biological sample from the subject; and
25 b. comparing said amount to the mean
 amount in a normal subject, wherein a
 significant deviation from said mean amount in
 a normal subject is indicative of an
 inflammatory pathology in said subject.

30

14. The method of claim 13, wherein the presence of said cryptdin is determined by contacting said biological sample with a detectable anti-cryptdin antibody and detecting binding of said anti-cryptdin
35 antibody to said biological sample.

15. The method of claim 13, wherein said biological sample is selected from the group consisting of intestinal tissue and the contents of the intestinal lumen.

5

16. The method of claim 13, wherein said inflammatory pathology is selected from the group consisting of inflammatory bowel disease, pancreatitis, malignancy, infection and ileitis.

10

17. A method for treating inflammation of the intestine in a subject, comprising administering a cryptdin peptide to the subject.

15

18. The method of claim 17, wherein said subject is immunocompromised.

19. The method of claim 18, wherein said subject is immunocompromised due to malignancy, malnutrition, radiation burns, immunosuppressive infections, autoimmune disease or neonatality, bone marrow transplantation or chemotherapy.

20. The method of claim 17, wherein said cryptdin is administered by a means selected from the group consisting of oral administration, nasogastric intubation, transabdominal catheterization, intravenous administration, aerosol inhalation and topical administration.

30

21. The method of claim 17, wherein more than one cryptdin is administered simultaneously or sequentially.

22. The method of claim 17, wherein said cryptdin is administered orally in a delayed release formulation designed to permit release in the small intestine.

5

23. An anti-cryptdin antibody.

24. The anti-cryptdin antibody of claim 23, wherein said antibody is a monoclonal antibody.

10

25. In a method for chemically synthesizing a peptide by attaching an amino acid to a resin, sequentially coupling additional amino acids to obtain a protected peptide resin, cleaving the protected peptide from the resin and deprotecting the peptide, the improvement comprising prior to cleavage and deprotection, reswelling the protected peptide resin with dichloromethane.

20

26. The method of claim 25, wherein said peptide is a cryptdin.

27. A method for preventing inflammation in a subject as a result of surgery, comprising administering a cryptdin to said subject prior to said surgery.

25

28. A substantially purified nucleic acid molecule encoding a cryptdin.

30

29. The nucleic acid molecule of claim 28, wherein said nucleic acid molecule is a cryptdin gene or a portion thereof.

30. A nucleic acid molecule, comprising a gene selected from the group consisting of the mouse cryptdins 1, 2, 3, 5, 6 and i genes as shown in Figure 11 (SEQ ID NOS: 53-58).

35

31. A nucleic acid molecule, comprising a gene selected from the group consisting of the rat cryptdins 1, 2 and 3 genes as shown in Figures 15.A. to 15.C. (SEQ ID NOS: 66-68).

5

32. The nucleic acid molecule of claim 28, wherein said nucleic acid molecule is a cryptdin cDNA sequence or a portion thereof.

10

33. A nucleic acid molecule, comprising a cDNA sequence selected from the group consisting of the mouse cryptdins 1-17 cDNA sequences as shown in Figure 10 (SEQ ID NOS: 34-50).

15

34. A nucleic acid molecule, comprising a cDNA sequence selected from the group consisting of the mouse cryptdins 2-17 cDNA sequences as shown in Figure 10 (SEQ ID NOS: 35-50).

20

35. A nucleic acid molecule, comprising a cDNA sequence selected from the group consisting of the rat cryptdins 1, 2 and 3 cDNA sequences as shown in Figures 14.A. to 14.C. (SEQ ID NOS: 63-65).

25

36. A nucleotide sequence that can hybridize under relatively stringent conditions to a nucleic acid molecule encoding a cryptdin.

30

37. The nucleotide sequence of claim 36, comprising a portion of a nucleic acid molecule selected from the group consisting of the mouse cryptdins 2-17 cDNA sequences as shown in Figure 10 (SEQ ID NOS: 34-50); the mouse cryptdins 1, 2, 3, 5, 6 and 1 genes as shown in Figure 11 (SEQ ID NOS: 53-58); the rat cryptdins 1, 2 and 3 genes as shown in Figures 15.A. to 15.C. (SEQ ID NOS: 66-68) and the rat cryptdins 1, 2 and 3 cDNA sequences as shown in Figures 14.A. to 14.C. (SEQ ID NOS: 63-65).

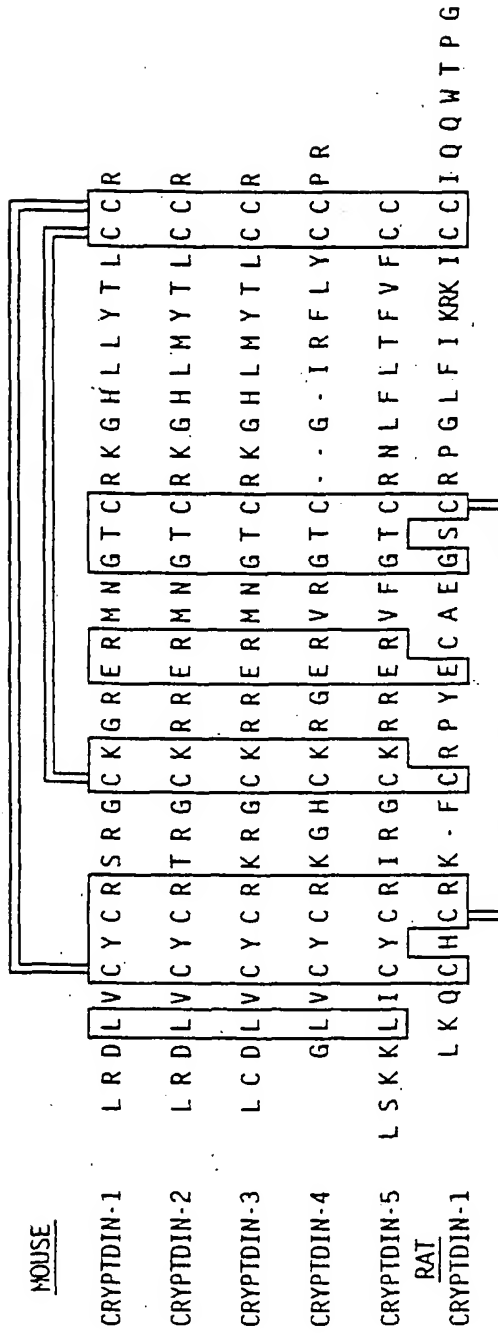
35

38. A method of detecting the presence of a nucleic acid molecule encoding a cryptdin in a biological sample, comprising the steps of:

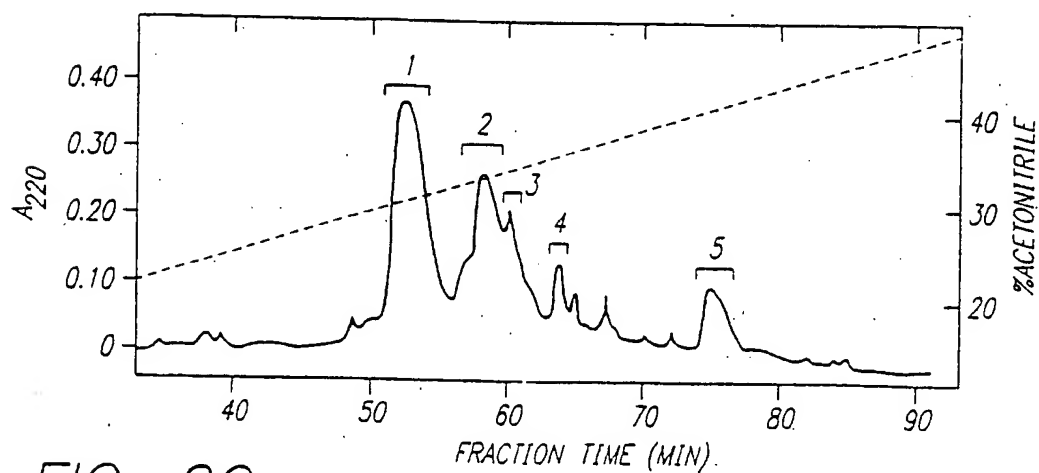
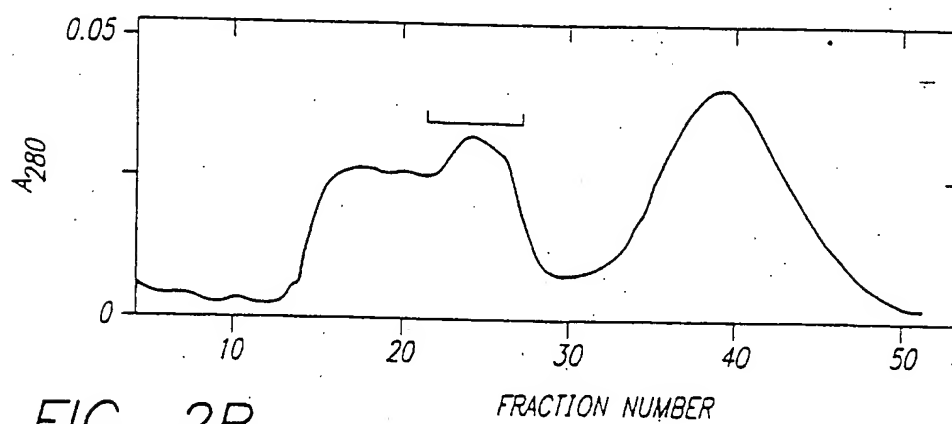
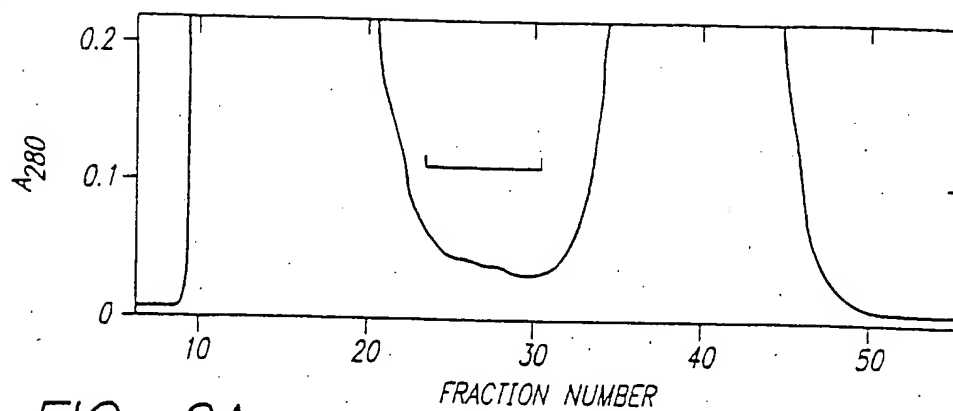
5 a. contacting the biological sample with the nucleotide sequence of claim 35 under relatively stringent hybridization conditions; and

10 b. detecting hybridization of said nucleotide sequence to a nucleic acid molecule present in said sample, wherein said hybridization indicates the presence of a nucleic acid molecule encoding a cryptdin.

FIG. 1



2/27



SUBSTITUTE SHEET (RULE 26)

FIG. 3

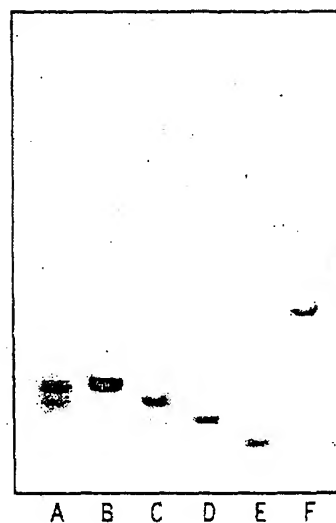
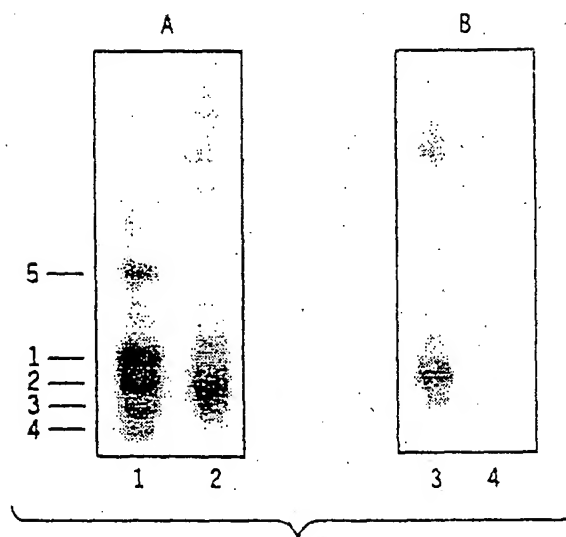


FIG. 4



SUBSTITUTE SHEET (RULE 26)

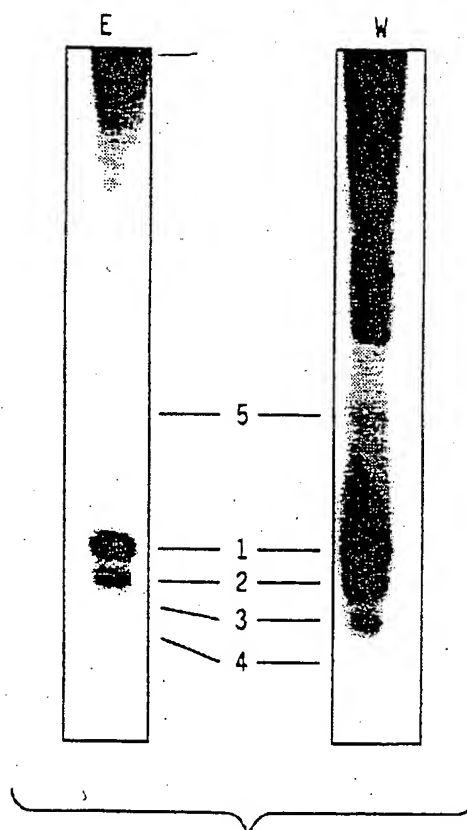


FIG. 5

SUBSTITUTE SHEET (RULE 26)

FIG. 6A

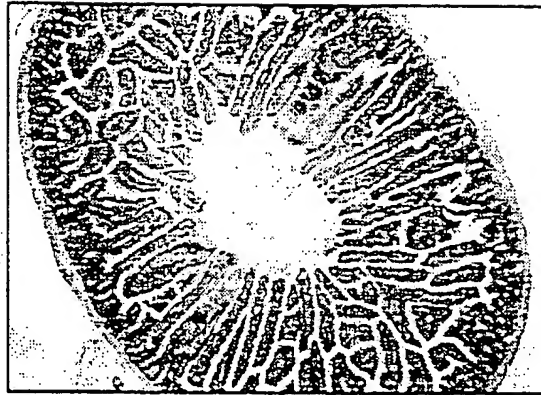
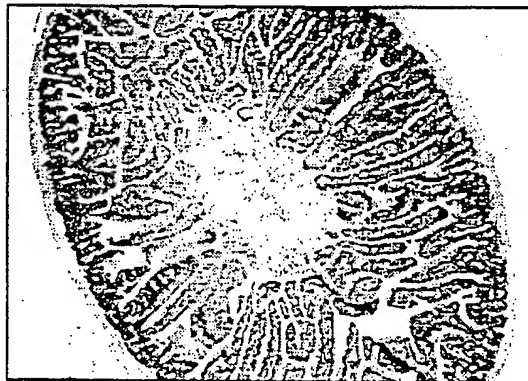


FIG. 6B



SUBSTITUTE SHEET (RULE 26)

FIG. 6C

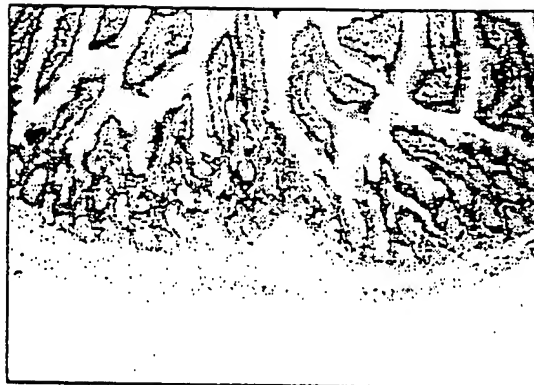
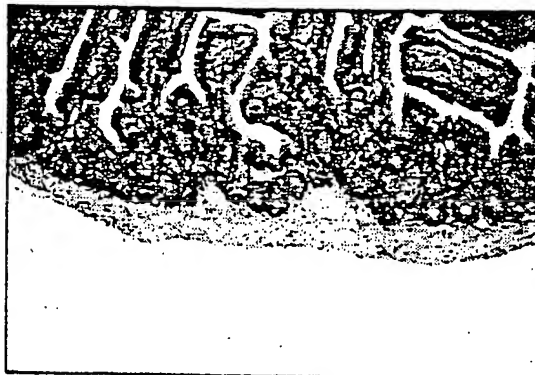


FIG. 6D



SUBSTITUTE SHEET (RULE 26)

FIG. 6E

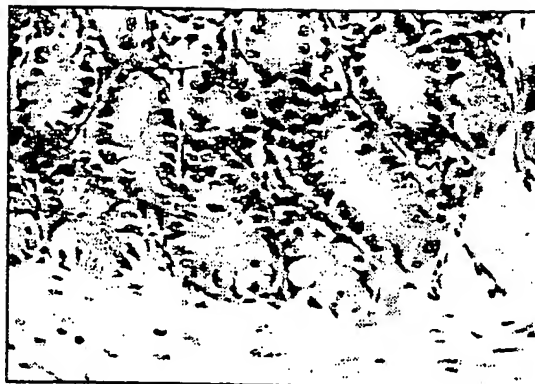


FIG. 6F



SUBSTITUTE SHEET (RULE 26)

8/27

FIG. 7A

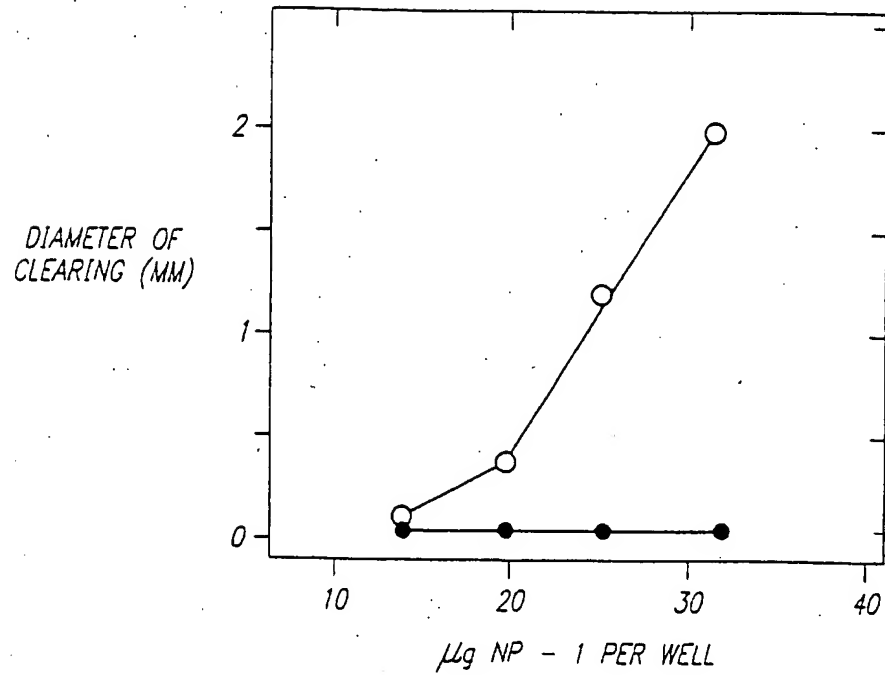
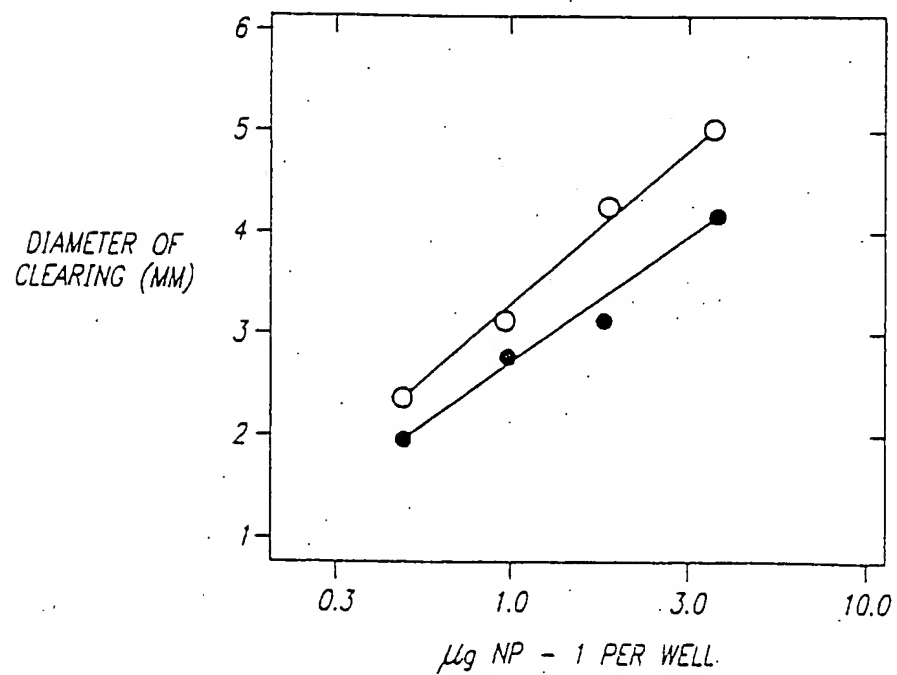


FIG. 7B



9/27

FIG. 8

RAT CRYPT-1 LKQCHCRKF CRPYEKAEGSCRPGLFIKRKICCIQQWTPGRT
 RAT CRYPT-2 IGRPVRRRCRAN CGPKEYATAFCAQGPFKQKFCCT
 RAT CRYPT-3 IRWPWKRCRCSF CRPYENATSFCAQGLFKQHKFCCLDTWPPRMK

HUMAN HD-5 TSGSQARATCYCRTGRCATRESLSGVCEISGRL YRLCCR
 HUMAN HD-6 AFTCHCRR - SCYSTEYSYGTCTVMGIN HRFCCCL

Def CONSENSUS

CxC R x x x C x x x E R x x G x C x x x x x x x x C C

RAT CRYPT-1 (cDNA)	1	25	50
RAT CRYPT-2 (genomic)	MKTLVLLSALVLLAFQVQADPIQEAEEEEETKTEEQPADEDQDVSVSFEGPE		
RAT CRYPT-3 (cDNA)	MKTLVLLSALVLLAFQVQADPIQGAEEEEETKTEEQPSDEDQDVSVSFEGPE		
	MKTLVLLSALVLLAFQVQADPIQEAEEEEETKTEEQPADEDQDVSVSFEGPE		
	51	75	100
RAT CRYPT-1	PSALQNLEIGWPLKQCHCRKFRCRPFYKAEAGSCRPGLFIKRKICCIQQWTPGRT		
RAT CRYPT-2	ASALQDFEIGRPVRRRCRANCGRPKYATAFCAQGPFKQKFCCT		
RAT CRYPT-3	PSALQNLEIRWPWKRCRCSFRCRPFYENATSFCAQGLFKQHKFCCLDTWPPRMK		

SUBSTITUTE SHEET (RULE 26)

10/27

FIG. 9A

	Exon 1	Exon 2
	<---SIGNAL PEPTIDE---><----->	----->
	PROPIECE	CRYPTIDIN
	1	58
Cryp01	MKTLVLLSAL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GTSLQEEES LRDLCVCYCRS RGCKGRERMIN GTCRKGHILLY TLCCR	
Cryp02	MKTLVLLSAL VLLAFQVQAD PIQNTDEETK TEKQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRT RGCKRRERMIN GTCRKGHILMY TLCCR	
Cryp03	MKTLVLLSAL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKRRERMIN GTCRKGHILMY TLCCR	
Cryp04	...LVLLSAL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK GHCKRGERVR GTC**G*IRF LYCCPRR	
Cryp05	MKTFVLLSAL VLLAFQVQAD PIHKTDEETN TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKRRERVF GTCRNLFLTF VFCCS	
Cryp06	MKTLVLLSAL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKGRERMIN GTCRKGHILLY MLCCR	
Cryp07	MKTLVLLSAL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRT RGCKRRERMIN GTCRKGHILMY TLCCR	
Cryp08	MKTLVLLSAL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKRRERMIN GTCRKGHILMY TLCCR	
Cryp09LAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKRRERMIN GTCRKGHILMY TLCCR	
Cryp10LAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKGRERMIN GTCRKGHILLY TLCCR	
Cryp11AL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRS RGCKGRERMIN GTCRKGHILLY MLCCR	
Cryp12	MKTLVLLSAL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKRRERMIN GTCRKGHILMY TLCCR	
Cryp13	MKTLVLLSAL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKRRERMIN GTCRKGHILMY TLCCR	
Cryp14AL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKRRERMIN GTCRKGHILMY TLCCR	
Cryp15	MKTLVLLSAL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKRRERMIN GTCRKGHILLY MLCCR	
Cryp16	MKTLVLLSAL VLLAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRS RGCKRRERMIN GTCRKGHILMY TLCCR	
Cryp17LAFQVQAD PIQNTDEETK TEEQPGEDDQ AVSVSFGDPE GSSLQEEES LRDLCVCYCRK RGCKRRERMIN GTCRKGHILLY TLCCR	

SUBSTITUTE SHEET (RULE 26)

11/27

FIG. 9B

Cryp01	MKTLVLLSAL	VLLAFQVQAD	PIQNTDEETK	TEEQPGEDDQ	AVSVSFGDPE	GTSIQEES	LRLVVCYCRS	RGCKGRMRN	GTCRKGHILLY	TLCCR
Cryp02	I			K	E	S	T	R		M
Cryp03						S	K	R		M
Cryp04					I	GQ	SA.H.K.	K.GH.RG	VR	**.*TRF LY..PRR
Cryp05	F		HK	N		I	GQ	SA.H.	LSKK.I	I
Cryp06					E		A			NLF.TF VF..S
Cryp07	I					S	T	R.H.		M
Cryp09					E	S	K	R.H.		M
Cryp08					E	S	K	R.H.		M
Cryp10						S	K	R.H.		M
Cryp11					E					M
Cryp12	I				E		A			M
Cryp13					E		K	R.H.		R.M.
Cryp14					E	S	T	R		M
Cryp15						S	K	R.HI		M
Cryp16	I				E					M
Cryp17					E	I	K	R.H.		

SUBSTITUTE SHEET (RULE 26)

FIG. 10(a)

[illegible]

13/27

FIG. 10(b)

Cryptdin-1	AspProIleGlnAsnThrAspGluGluThrLysThrGluGluGlnProGlyGluAspAspGlnAlaValSerPheGlyAspProGluGlyThr	20	30	40	50						
Codon:											
Cryp01	CTGATCCTATCCAAACACAGATGAAGAGACTAAAACTGAGGAGCAGCCAGGGGAAGACGACCGGCGTATCTGTCTCCTTTGGAGAGCCAGAAAGGCAC	60	70	80	90	100	110	120	130	140	150
Base											
Cryp01
Cryp02
Cryp03
Cryp04
Cryp05
Cryp06
Cryp07
Cryp08
Cryp09
Cryp10
Cryp11
Cryp12
Cryp13
Cryp14
Cryp15
Cryp16
Cryp17
Consensus	CTGATCCTATCCAAACACAGATGAAGAGACTAAAACTGAGGAGCAGCCAGGGGAAGA-GACCAGGCTGTGTCGTCTCTTTGGAGAGCCAGAAAGGC-C										

SUBSTITUTE SHEET (RULE 26)

14/27

FIG. 10(c)

Cryptdin-1	SerLeuGlnGluSerLeuArgAspLeuValCysTyrCysArgSerArgGlyCysLysGlyArgGluArgMetAsnGlyThrCysArgLysGlyHis	60	70	80
Codon:				
Cryp01	TTCTCTTCAGAGGAATCGTTGAGAGAICTGGTATGCTATTGTAGATCAAGAGGCTGCAAGAGGAAGAGAACGCATGAATGGGACCTGCAGAAAGGGTCAT	160	170	180
Base				
Cryp01TC.....G.....G.....A.....	190	200	210
Cryp02TC.....A.....G.....	220	230	240
Cryp03AA.....A.....G.....	250		
Cryp04	G.....T..AA...T.....G.T..T.....AA.G.CA.....A.G.....GAG.TCG.....T.TG...TAC.ATT.			
Cryp05	.G.....T.....T.CA.A.A.G..A.....AT.....A.....A.....G.G.TTT.....TCT.TT.			
Cryp06A.....T.....GC.....G.....G.....			
Cryp07AC.....A.....A.....			
Cryp08AA.....A.....A.....			
Cryp09AA.....A.....A.....			
Cryp10AA.....AA.....G.....G.....A.....			
Cryp11TC.....TC.....G.....G.....A.....			
Cryp12C.....A.....GC.....G.....G.....			
Cryp13AA.....AA.....A.....A.....G.....			
Cryp14AC.....AC.....A.....G.....			
Cryp15AA.....AA.....A.....A.....A.....			
Cryp16TC.....TC.....G.....G.....			
Cryp17T.....T.....AA.....A.....A.....			
Consensus	TTCTCTTCAGAGGAATCGTTGAGAGAICTGGTATGCTATTGTAGA--AAGAGGCTGC AAA-GAAGAGAAC-CATGAATGGGACCTGCAGAAAGGGTCAT			

SUBSTITUTE SHEET (RULE 26)

15/27

FIG. 10(d)

Cryptdin-1 LeuLeuTyrThrLeuCysCysArgEnd																
Codon:	***															
Cryp01	TTATTGTACACGGCTCTGCTGCTGAacatggagaccacagaggacgaacatgagtactgagccactgatgctggtgcctgatgaccacttcg															
Base:	260	270	280	290	300	310	320	330	340	350						
Cryp01	T	C		A	G	CG	A			C	G					
Cryp02	A	C		A	G	TG	C									
Cryp03	A	CA		A	G	CG	A									
Cryp04	GTAC.G.TGC.C.C..C.CT.AAC.TGCA.AT..AAAGATATGAC.ACCATTGTCTCTG.GGCC.CTG.TGCCG.GGCC..ATGACCACTTCTCAA.															
Cryp05	ACT.T.GTAT		CA		T	C	TGA		A	ATATGACAACCATC.GC.CTGAG.C	ACTGATGCTG		GT	TGATGAT	AC	CGC
Cryp06	T	T		T	A	CG	C									
Cryp07	A	C		T	A	CG	C									
Cryp08	A	C		A	G	CA	G									
Cryp09	T	T		A	G	TG	A									
Cryp10	T	C		A	G	CG	A									
Cryp11	T	T		A	A	CG	C									
Cryp12	A	C		T	A	CG	C									
Cryp13	A	CA		A	G	CG	A					C	G			
Cryp14	A	C		A	G	CA	G									
Cryp15	T	T		A	G	TG	C					CG*				
Cryp16	A	C		T	A	CG	C									
Cryp17	T		T	A	G	T	C					C	G			
Consensus	TTA-TGTACA-GCTCTGCTGCTGCTGAacatggagacc-cagag-acaaga-a-qatgagtactgagccactgatgctggtgcctgatgaccacttct															

SUBSTITUTE SHEET (RULE 26)

16/27

FIG. 10(e)

Cryp01	caataaattgttcgcaatatgc(A) _n
Base	360 370
Cryp01
Cryp02C.T.....
Cryp03
Cryp04	A...-GT...CAATATGC.
Cryp05	A.TA..T.GT.CGCA.TATG.
Cryp06C.....
Cryp07
Cryp08
Cryp09
Cryp10
Cryp11C.....
Cryp12
Cryp13
Cryp14A.....
Cryp15
Cryp16
Cryp17C.T.....
Consensus	caataaattgttcgcaatatgc

17/27

FIG. 17(a)

Nucleotide	CODON									
	1	2	3	4	5	6	7	8	9	
tataaatgcaxxctggxtxtcactxtccacacattgxcgtcctgctcaccaatcctccaggtagctcccgagccatgaagacacttgctcctctctcgc	1	2	3	4	5	6	7	8	9	
.....gg.....a.at.....g.....A.....	1	2	3	4	5	6	7	8	9	
.....ga.....c.cc.....C.....G.....CA.....	1	2	3	4	5	6	7	8	9	
.....gg.....a.at.....c.....G.....	1	2	3	4	5	6	7	8	9	
.....agt.....c.ac.....c.....G.....T.....	1	2	3	4	5	6	7	8	9	
.....gg.....a.at.....c.....C.A.....G.A.....AA.....	1	2	3	4	5	6	7	8	9	
ac.tg.g.gtaa.a.cc.tc.caat.....A.....	1	2	3	4	5	6	7	8	9	
Cryp1	30	40	50	60	70	80	90	100	110	120
Cryp2	10	11	12	13	14	15	16	17	18	19
Cryp3	20	21	22	23	24	25	26	27	28	29
Cryp5	30	31	32	33	34	35	36	37	38	39
Cryp6	40	41	42	43	44	45	46	47	48	49
Cryp1	50	51	52	53	54	55	56	57	58	59
Cryp2	60	61	62	63	64	65	66	67	68	69
Cryp3	70	71	72	73	74	75	76	77	78	79
Cryp5	80	81	82	83	84	85	86	87	88	89
Cryp6	90	91	92	93	94	95	96	97	98	99
Cryp1	100	101	102	103	104	105	106	107	108	109

SUBSTITUTE SHEET (RULE 26)

8NSDOCIO: <WO 9616075A1_1_>

SUBSTITUTE SHEET (RULE 26)

[illegible]

CODON	91	92	93	280
	TGCTGXGCTGACATGGAGA			
Cryp1TC.....			
Cryp2CC.....			
Cryp3TC.....			
Cryp5TA.....T...C...			
Cryp6TC.....			
Cryp1CT.....			

FIG. 11(b)

19/27

FIG. 12A

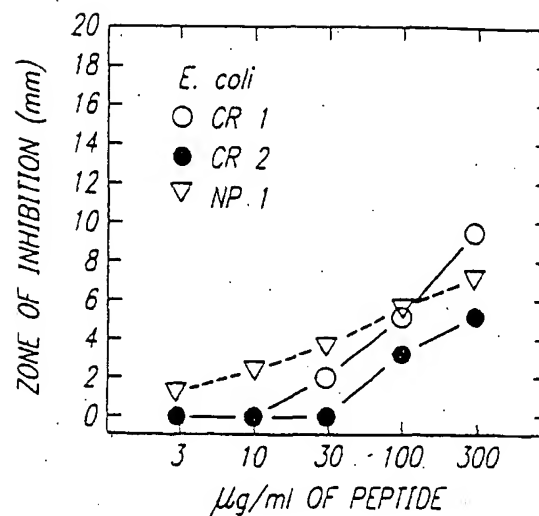


FIG. 12B

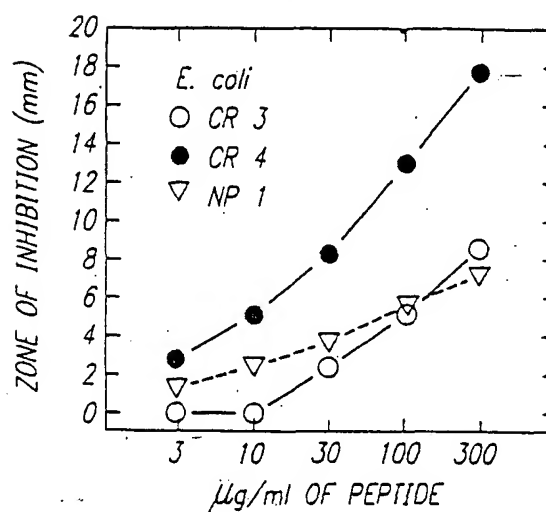
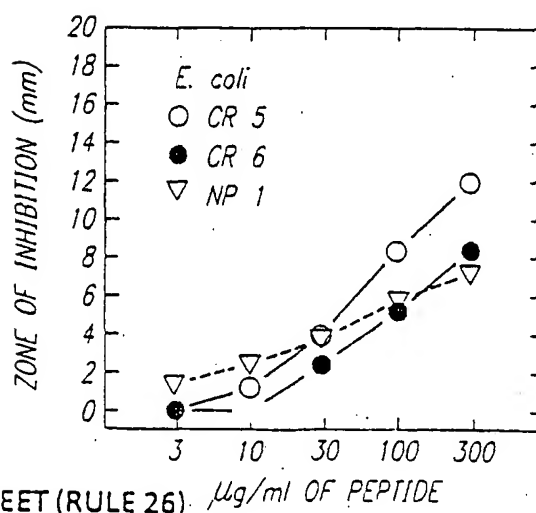


FIG. 12C



SUBSTITUTE SHEET (RULE 26) µg/ml OF PEPTIDE

20/27

FIG. 13A

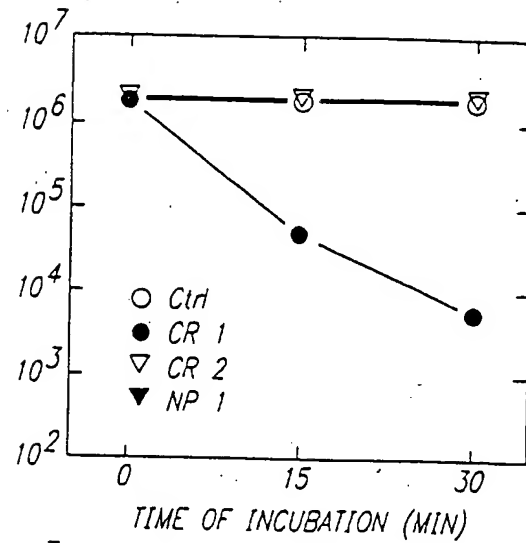


FIG. 13B

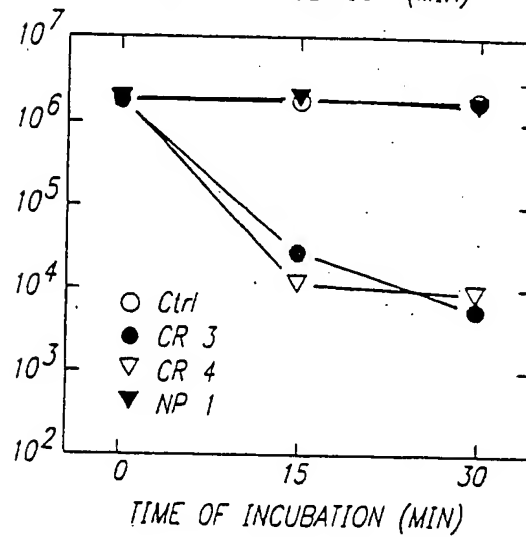
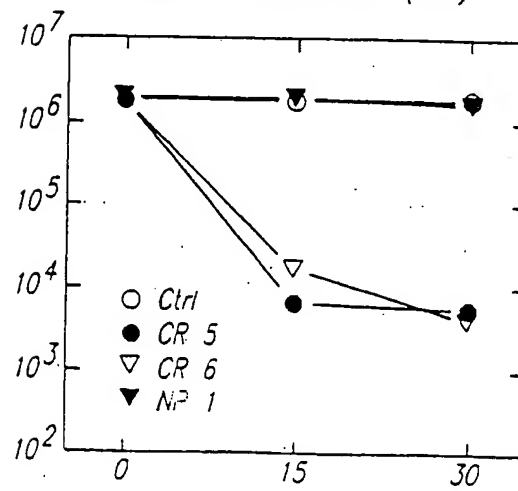


FIG. 13C



SUBSTITUTE SHEET (RULE 26) TIME OF INCUBATION (MIN)

21/27

FIG. 14A

RAT CRYPTIN 1 cDNA SEQUENCE

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
ACACTGGTCT	CCAGCTCACC	AATCCTCCAG	GTGACTTCCA	GCCATGAAGA	50
CTCTTGTCT	CCTCTCTGCC	CTTGCTCTGC	TGGCATTCCA	GGTCCAGGCT	100
GATCCCATT	AAGAGGCAGA	AGAAGAGACT	AAAAGTGAAG	AGCAGCCAGC	150
AGATGAGGAC	CAGGATGTGT	CTGTCTCCTT	TGAAGGCCCA	GAACCTCTG	200
CTCTTCAAAA	TTTAGAGATA	GGATGGCCAT	TAAAGCAGTG	CCATTGCCGA	250
AAGTTCTGCA	GACCTTATGA	AAAGGCCGAG	GGGTCTGTG	GTCCAGGTCT	300
ATTTATAAAA	CGCAAAATCT	GCTGCATACA	ACAATGGACA	CCAGGGAGGA	350
CATAACCACG	TGAACTGGGA	CCTCACAATC	TGTCATTCTT	GGGCTTCAAC	400
TCGAETGCTT	TTCCTTCTCC	AATAAACCCC	TTGCAGACAA	AAAAA	445

FIG. 14B

RAT CRYPTIN 2 cDNA SEQUENCE

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
ACACTGGTCT	CCAGCTCACC	AATCCTCCAG	GTGACTTCCA	GCCATGAAGA	50
CTCTTGTCT	CCTCTCTGCC	CTTGCTCTGC	TGGCTACCA	GGTCCAGGCT	100
GATCCCATT	AAGGGGCAGA	AGAAGAGACT	AAAAGTGAAG	AGCAACCATC	150
AGATGAGGAC	CAGGATGTGT	CTGTCTCCTT	TGAAGGCCCA	GAAGCTCTG	200
CTCTTCAAGA	TTTTGAGATA	GGAAGGCCAG	TGAGGAGGTG	CCGTTGCAGA	250
GCAAACTGCG	GACCTAAAGA	ATATGCCACT	GCGTTCTGTG	CTCAAGGTCC	300
ATTTAAACAG	TTCAAATTCT	GCTGCACATG	AACATGGATC	CCAAGTCTGA	350
GATAACCACG	TGCTCTGGGA	CCTCACAATC	TGTCATTATT	GTGCTTGACC	400
TCAACTGCTT	TTCCTTCTCC	AATAAACTCC	TGGCAGACAA	AAAAA	445

FIG. 14C

RAT CRYPTIN 3 cDNA SEQUENCE

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
ACACTGGTCT	CCAGCTCACC	AATCCTCCAG	GTGACTTCCA	GCCATGAAGA	50
CTCTTGTCT	CCTCTCTGCC	CTTGCTCTGC	TGGCATTCCA	GATCCAGGCT	100
GATCCCATT	AAGAGGCAGA	AGAAGAGACT	AAAAGTGAAG	AGCAGCCAGC	150
AGATGAGGAC	CAGGATGTGT	CTGTCTCCTT	TGAAGGCCCA	GAACCTCTG	200
CTCTTCAAAA	TTTAGAGATC	AGATGGCCAT	GGAAGAGGTG	CCATTGCAGA	250
AGTTTCTGCA	GACCTTATGA	AAATGCCACT	TCGTTCTGTG	CTCAAGGTCT	300
ATTTAAACAA	CACAAATTCT	GCTGCCTAGA	AACATGGCCC	CCAAGGATGA	350
AATAACCACG	TGCTCTGGGA	CCTCACAATC	TGTCATCATT	GTGCTTGGCC	400
TCAACTTCTT	TTCCTTCTCC	AATAAACTCC	TTGCAGACAA	AAAAA	445

SUBSTITUTE SHEET (RULE 26)

22/27

FIG. 15A(1)

RAT CRYPTIN 1 GENOMIC SEQUENCE

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
CCTGAGACCA	ACTCTGTGAT	AATCAGAAAA	GTCAATAATG	TGTCTGAAAT	50
GTAAGGTGTG	CTTCTTGACT	GATAGTTCTA	AGCCTACAGA	GAGATTCATG	100
TGGTCATATC	CCATTTAACA	ATGATATATA	TGTTAAATAT	ATAAAGATAT	150
ATGTATGTTT	AGTATGTATG	TTCAATATGT	ATGTAAATAA	TATTCTTGCT	200
GCTTCACTAG	CTTTTACACA	GAGCTGTAAG	TAAAAACATT	GTAGCCAATG	250
AATAGTATTT	ATTAACATGT	AAATAGGAGC	TGGCACCTGT	GACAGTGGGA	300
CTCCATACAC	TGACTGTAAA	CAACAGGATG	CTCTGGACCT	TTTGCTGTGT	350
GTGTGGTGAG	AGACATGGGA	TAAACACAGA	CTGAAGAGTG	TTCTGAATG	400
ACATGGCGGC	ACTTCTCGAG	ACCGGGTAGC	AGCTTCTGAG	CCTCTCTACA	450
TTGTGGATGT	CCTTTCCTGT	AGGTCAGGTC	TCATTGTCTA	AAAGTAAAG	500
CATTGCAGCA	TCTCAGACCT	GGGAAACACC	CCATGGCTTG	AGGGTCTCTGA	550
GCATGAAGAG	CCACCTGGAG	CTCACTCTTG	GCAGATGTGT	TCCATGACTT	600
TGGCTTCTTC	AGAACAACCC	ACTACAGCTT	CACTCTGACA	AATCCTAGAA	650
ACTTGAACCT	AATTCACCTAG	AGGGCACCAT	AAAGCCATCA	TACCTTATAA	700
TGGCCCCAAA	GGAGGTGATT	CACAAAGTTT	GCCTTGATGA	GGACAATTGC	750
TAATACACAA	AACTTGCAA	AAAAAATTG	AGTGTCCAGT	CCACCTGGTC	800
AAGGACTGGT	CCCGGATCCA	CAGTTTCTGA	GAATAGCAGG	CTCTAACTTG	850
AAAACACAAA	AATTGTTTGT	TCTATGAGCT	CATTAAATTA	GGCAGTGTTT	900
AGCTATTTTC	TTTCTGACC	ACTGAGAGGT	AAATACTCAA	GCAGATGGGA	950
AACAGGGGAG	GACAGTAAAG	CCTGTTTCATC	ATTATCAGTG	GGAGTGTGCA	1000
TGAGGGGAGG	GGTGTCAAGT	AACACACAGA	GCATCAGGAA	GGAAGCCTTG	1050
AGGACAGAGG	AACATCAAAG	GGATCCTGAG	GACAACAGCT	GGGAGCAGTT	1100
GCCATCAATG	AGTGCCTTCT	CTAAGTATGG	GGCATGTTCT	TTGCCCTATA	1150
AATGCAGGCT	GGCTTCTCTC	TCCACACACT	GGTCTCCAGC	TCACCAATCC	1200
TCCAGGTGAC	TTCCAGCCAT	GAAGACTCTT	GTCTCTCTCT	CTGCCCTTGT	1250
CCTGCTGGCA	TTCCAGGTCC	AGGCTGATCC	CATTCAAGAG	GCAGAAGAAG	1300
AGACTAAAC	TGAGGAGCAG	CCAGCAGATG	AGGACCAGGA	TGTGTCTGTC	1350
TCCTTTGAAG	GCCCAGAACC	CTCTGCTCTT	CAAAATTTAG	GTGCGTGCTT	1400
GTGCACAGAA	TGATGGAGGC	TTGGAGTCTC	CTGATGGAGG	GTTGTAGATT	1450
AGCCCTGGAG	TCCTGTCAAG	GACAGTCTGG	TTGAGGTAGC	TGTCTACTGA	1500
TCCTTTTCAGA	ACTTCCCTGT	CTTATTCATA	GAAATAACAG	TGAGAGACAA	1550
GCCATTGGGC	TTGACTTTTT	CCTTTTAAGA	TTTCGGTCTA	ACAATTTATC	1600
TGTGAAAAAC	CTTTAAAATA	TAAAACATAT	TGATTAGTTC	TTTAAACCTG	1650
AGTGATAATT	TTCTTACAGG	AAGAAATATC	CGTTTTACCC	TAAAAATTAG	1700
ATTGGTACCC	AAATGCCAGT	GTATGAAGGT	GTTGGGTCAA	GAAAACACAA	1750
AAAACTGTT	AGAATATGGT	GATAGTAAAA	ATTCCTATAT	GTGATTAACA	1800
CTTGTTAAAC	ATCTTATCTC	CATGTGTTTG	GGGTGATCA	CTGTGCTGGC	1850
TGTGATGTCA	CCCACACAGC	AAACCTACTC	TCTACCATGC	ACAGGACATC	1900

SUBSTITUTE SHEET (RULE 26)

23/27

FIG. 15A(2)

RAT CRYPTIN 1 GENOMIC SEQUENCE

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
TTCATGGGGT	AGTTCACTGT	TACACACTAC	TGGCCTCCTT	ACTTCATGCC	1950
TGATGCTTTC	TTGTTTCCTC	AGAGATAGGA	TGGCCATTAA	AGCAGTGCCA	2000
TTGCCGAAAG	TTCTGCAGAC	CTTATGAAAA	GGCCGAGGGG	TCCTGTCGTC	2050
CAGGTCTATT	TATAAAACGC	AAAATCTGCT	GCATACAACA	ATGGACACCA	2100
GGGAGGACAT	AACCACGTGA	ACTGGGACCT	CACAATCTGT	CATTCTTGGG	2150
CTTCAACTCG	ACTGCTTTTC	CTTCTCCAAT	AAACCCCTTG	CAGACAAATA	2200
ACCTGTTTAT	GTTTTTTTGA	TGCTTTCTAT	GTGGCGTAGA	CAGGACTCTC	2250
CTGAGCCATG	TAGCAAAATC	TTCAGTGAAT	CCTTTGTAAA	AGAAGTCTTG	2300
GTCACATTTC	AGCAGTCATA	TCAAGGATGA	GCAGGAGGTT	AGATCCAAAG	2350
AGACAAGATG	GTCTGCGCCA	GCTGCTTCTG	TGTCTATCAA	GTCTTCTGTC	2400
CTTTAGATTA	GAGTCACCCCT	CAAAAATTAG	TTCCAGATTT	TCATGTTCTA	2450
TTTTTTC					2457

SUBSTITUTE SHEET (RULE 26)

24/27

FIG. 15B(1)

RAT CRYPTIN 2 GENOMIC SEQUENCE

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
TATTACGAAT	TCGAGCTCGG	TACCGGTATA	TGAAGAGCGA	CCACTGCCAG	50
GACGAAAGTG	CAATGCCGCA	TACCTCAGTG	GCGTGGAGTG	CAGGTATACA	100
GATTAATCCG	GCAGCGTCCG	TCGTTGTTGA	TATTGCTTAT	GAAGGCTCCG	150
GCAGTGGCGA	CTGGCGTACT	GACGGATTCA	TCGTTGGGGT	CGGTTATAAA	200
TTCTGATTAG	CCAGGTAACA	CAGTGTATG	ACAGCCCGCC	GGAAACCGGTG	250
GGCTTTTTTG	TGGGGTGAAT	ATGGCAGTAA	AGATTTCAGG	AGTCTGAAA	300
GACGGCACAG	GAAAACCGGT	ACAGAACTGC	ACCATTTCAGC	TGAAAGCCAG	350
ACGTAACAGC	ACCACGGTGG	TGGTGAACAC	GGTGGGCTCA	GAGAATCCGG	400
ATGAAGCCTG	CTTTTTTATA	CTAAGTTGGC	ATTATAAAAA	AGCATTGCTT	450
ATCAATTTGT	TGCAACGAAC	AGGTCACAT	CAGTCAAAAT	AAAATCATT	500
TTTGATTTC	ATTTTGTCCC	ACTCCCTGCC	TCTGTCATCA	CGATACTGTG	550
ATGCCATGGT	GTCCGACTTA	TGCCCGAGAA	GATGTTGAGC	AACTTATCG	600
CTTATCTGCT	TCTCATAGAG	TCTTGCAGAC	AACTGCGCA	ACTCGTGA	650
GGTAGGCGGA	TCTGGGTGCA	CTTAGGCCT	CACTGGCCTA	ATACGACTCA	700
CTATAGGGAG	CTCGAGGATC	ATTGCTAATA	CCATGAAACT	TGACCACCTG	750
GTCAAGGACT	GGTCCAGGGT	CCACAGTTTC	TGAGAAGAGC	AGGCTCCAAC	800
TTCTAACCAC	AAAACTATT	TTTTCCATGC	GCTCCTTAAA	TTAGGCAGCG	850
CCCAGCTATT	TTCTTTCTG	ACCACTGAGA	GGTAAATACT	CAAGCAGATG	900
GGAAACAGGG	GAAGATAGCA	AGGCCTCTTC	ATCATTATCA	CTGGGTGTGT	950
GCGTGAGGGG	AGGGGTGTCA	TTGCATACAC	AGGGCAACAT	CAGGATGGAA	1000
GCCTTGAGGA	CAGAGGAACA	TCAAAGGGAT	CCTGAGGACA	ACAGCTGGGA	1050
GCAGTTGCCA	TCAGTGAGTG	CCTTCTCTAA	GTGTGGGGCC	TTTCTCTGCC	1100
ACATAAATGC	AGGCTGCCTC	CTCTCTCCAC	ACACTGGTCT	CCAGCTCACC	1150
AATCCTCCAG	GTGACTTCCA	GCCATGAAGA	CTCTTGTCTT	CCTCTCTGCC	1200
CTTGTCTCTG	TGGCCTACCA	GGTCCAGGCT	GATCCCATTC	AAGGGGCAGA	1250
AGAAGAGACT	AAAACTGAAG	AGCAACCATC	AGATGAGGAC	CAGGATGTGT	1300
CTGTCTCCTT	TGAAGGCCCA	GAAGCCTCTG	CTCTTCAAGA	TTTTGGTGAG	1350
TGCTTATGCA	CAGAATGATG	GAGGCTTGA	GTCTCCTGAT	GGAGGGTTGT	1400
AGATTAGACC	TGGAATCCTG	TCAAGAACTG	TCTGGTTCAG	GTAGCTGTCT	1450
CTTGGTCCCT	TTACATTCTT	TGTCTTCTTC	ATAGAAGTAA	CGGAGAGAGA	1500
TTAACCATTG	GGCTTGACTT	TTTTCTTTT	AAAATTTTTG	ATCTAACAAT	1550
TTATCTGTGG	AAAACCTTTA	AAATATAAAA	CATATTGATT	AGTTCTTTTA	1600
GACCTGATTG	ATAATTTTGT	TATAAGAAGA	AATATTCGTT	CTACTTTAAA	1650
AATTAGATTT	GGGACCCAAA	TGCCAGTGTA	TGAAGCTGTT	GGGTAAGGAA	1700
AAACCAAAAA	TGGTGATAGA	ATGTTGTGTA	GATGACAATT	CCTTTATGCG	1750
ATTAACACTT	TTTAAATGT	CTTATCTCCA	TGTGTTTGGG	GTTGATCATG	1800

SUBSTITUTE SHEET (RULE 26)

25/27

FIG. 15B(2)

RAT CRYPTIN 2 GENOMIC SEQUENCE

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
GTGCTGACTG	TGATGTCACC	CACAGAGCAA	ACCTACTCTC	TACCATGCAC	1850
AGGACATCTT	CATAGGGTAG	TTCAGTGCA	CACACTGCTG	GCCTCGTTAC	1900
TTCATGCCTG	ATGCTTTCTT	GTTTCCTCAG	AGATAGGAAG	GCCAGTGAGG	1950
AGGTGCCGTT	GCAGAGCAAA	CTGCGGACCT	AAAGAATATG	CCACTGCGTT	2000
CTGTGCTCAA	GGTCCATTTA	AACAGTTCAA	ATTCTGCTGC	ACATGAACAT	2050
GGATCCCAAG	TCTGAGATAA	CCACGTGCTC	TGGGACCTCA	CAATCTGTCA	2100
TTATTGTGCT	TGACCTCAAC	TGCTTTTCCT	TCTCCAATAA	ACTCCTGGCA	2150
GACAAATAAT	CGGTATATGT	TTATTTGATG	CTTTCTATTT	GGCTTAGACA	2200
GAAGCTCTCT	GAGCCATGTA	GCTGAATCTT	CAGTGAATCC	TTTGTAAGG	2250
TCACATTTC	GCAGTCATAT	CAAGGATGAG	CAGGAGGTTA	GATACAAAGA	2300
GACAAGATGG	TCTGCGCCAG	CTGCTTCTTT	GTCTATCAAG	TCTGCTTTCC	2350
TTTAGATTAG	AGTCACCATC	AAAAATTATT	CCCACATTTT	CATGTTCTAT	2400
ATTTTTTT					2408

SUBSTITUTE SHEET (RULE 26)

26/27

FIG. 15C(1)

RAT CRYPTIN 3 GENOMIC SEQUENCE

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
CCTGAGACCA	ACTCTGTGAT	AATCAGAAAA	GACAATTATG	TGTCTTAAAT	50
GTAAGGTTTG	CTTCTTGACT	GATAGATCTA	ACCCTACAGA	GAGATTCAAG	100
TGGTCTTGTC	CCATTGAACA	ATAGTATATA	TGTTTTATAT	ATATATATAT	150
ATATATGTAT	ATGTATATAT	ATATGTGTGT	GTGTGTGTGT	GTGTGTCTGT	200
GTCTGTGTGT	CTGTGTGTCT	GTGTGTCTGT	GTGTCTGTGT	GTGTATGTGT	250
GTGTATGTGT	ACATATGTTT	AATATGTCTG	TAAAATAGTA	TTCTTGTAGC	300
TTCACTTACT	TTTGACACAGA	GCTGTAAATA	AGAACATTGT	AGCCAATGAA	350
TAGTATTTAT	TAACATGTAA	ATAGGAGCTG	GCACCTCTGA	CAGTGGGACT	400
CCATACAGTG	ACTGTAAACA	ACAGGATGCT	CTAGACCTTT	TGCTGTGTGT	450
GTGGTGAGAG	ACATGGGATA	AACACAGACT	GAAGTGATG	ACATGGCGGC	500
ACTTCTCGAG	ACCGGGTAGC	AGCTTCTGAG	CCTCTCTACA	TTGTGGATGT	550
CCTTTCCTGT	AGTCCAGGTC	TCATTGTCTA	AAAGTAAAG	CATTGCAGCA	600
TCTCAGACCT	GGGAAACACC	CCATGGCTTG	AGGGTCCCGC	AGGTGAAGAG	650
CCACCTGGAG	CTCACTCTTG	GCAGATGTGT	TCCATGACTT	TGGCTTCTTC	700
AGAACCACCC	ACTACAGCTT	CACTCTGACA	AATCTTAGAA	ACTTGAAGTC	750
AATCACTGG	AGGGCACAAT	AAAGCCATCT	TACTTTCTCT	AAAATGGCCC	800
CAAAGGAGGG	GATTCACAAA	GTTTGCCTTG	ATGAGGACCA	TTGCTAATAC	850
CCCAAACTT	GCAAAAAAAA	TTGAGTGTCC	AGTCAACCTG	GTCAAGGACT	900
GGTCTGGAT	CCACAGTTTC	TGAGAAAAGA	AGGCTCCAAC	TTCAAAACAC	950
AAACCACTCC	TGTTCTATGC	GCTCATTAAA	TTAGGCAGTG	TTAAGCTATT	1000
TTCTTTCCTG	ACCACTGAGA	GGTAAATACT	CAAGCAGATG	GGAAACAGGG	1050
GAGGACAGCA	AAGCCTGTTC	ATCATTATCA	GTGGGAGTGT	GCGTGAGGGG	1100
AGGGGTGTCA	GTGAACACAC	AGAGCATCAG	GAAGGAAGCC	TTGAGGACAG	1150
AGGAACATCA	AAGGGATCCT	GAGGACAACA	GCTGGGAGCA	GTTGGCATCA	1200
GTGAGTGCCG	TCTCTAAGTG	TGGGGCCTTT	CTCTGCCACA	TAAATGCAGG	1250
CTGGCTCCTC	TCTCCACACA	CTGGTCTCCA	GCTCACCAAT	CCTCCAGGTG	1300
ACTTCCAGCC	ATGAAGACTC	TTGTCCTCCT	CTCTGCCCTT	GTCCTGCTGG	1350
CATTCCAGAT	CCAGGCTGAT	CCCATTCAAG	AGGCAGAAGA	AGAGACTAAA	1400
ACTGAGGAGC	AGCCAGCAGA	TGAGGACCAG	GATGTGTCTG	TCTCCTTTGA	1450
AGGCCCAGAA	CCCTCTGCTC	TTCAAAATTT	AGGTGCGTGC	TTGTGCACAG	1500
AATGATGGAG	GCTTGGAGTC	TCCTGATGGA	GGGTTGTAGA	TTAGCCCTGG	1550
AGTCCTGTCA	AGGACAGTCT	GGTTCAGGTA	GCTGTCTATT	GATCCTTTCA	1600
GAACCTCCCT	GTCTTATTCA	TAGAAATAAC	AGTGAGAGAC	AAGCCATTGG	1650
GCTTGACTTT	TTCCTTTTAA	GATTTTGGTC	TAACAATTTA	TCTGTGAAAA	1700
ACCTTTAAAA	TATAAAACAT	ATTGATTAGT	TCTTTTAAAC	CTGATTGATA	1750
ATTTTGTTAT	AGGAAGAAAT	AACTGTTCTA	CTTTAAAAAT	TAGATTTGGT	1800
ACCTAAATGC	CAGTGTATTA	AGGTGTTGGG	TCAGGAAAAC	ACAATAATGC	1850
TGATAGAATG	TGGTGTAGAT	GACAATTCCT	ATATGCGATT	AACACTTGTT	1900

SUBSTITUTE SHEET (RULE 26)

27/27

FIG. 15C(2)

RAT CRYPTIN 3 GENOMIC SEQUENCE

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
AAATTGTCCT	ATCTCCATGT	GTTTGGGGTT	GATCATGGTG	CTGGCTGTGA	1950
TGTCACCCAC	ACAGCAAACC	TACTTTCTAC	CATGCACAGG	ACATCTTCAT	2000
AGGGTAGTTC	ACTGTCACAC	ACTGCTGGCC	TCCTTACTTC	ATGCCTGATG	2050
CTTTCTCGTT	TCCTCAGAGA	TCAGATGGCC	ATGGAAGAGG	TGCCATTGCA	2100
GAAGTTTCTG	CAGACCTTAT	GAAAATGCCA	CTTCGTTCTG	TGCTCAAGGT	2150
CTATTTAAAC	AACACAAATT	CTGCTGCCTA	GAAACATGGC	CCCCAAGGAT	2200
GAAATAACCA	CGTGCTCTGG	GACCTCACAA	TCTGTCATCA	TTGTGCTTGG	2250
CCTCAACTTC	TTTTCTTCT	CCAATAAACT	CCTTGCAGAC	AAATAACCTG	2300
TTTATGTTTT	TTTGATGCTT	TCTATGTGGC	TTAGACAGGG	CTCTCCTGAG	2350
CCATGTAGCA	GAATCTTCAG	TGAATCCTTT	GTAAAAGAAG	TCTTGGTCAC	2400
ATTTCAACAG	TCATATCAAG	GATGAGCAGG	AGGTTAGATC	CAAAGAGACA	2450
AGATGCTCTG	CTCCAGCTGC	TTCTTGACTA	TCAAGTCTTC	TGTCCTTCAG	2500
ATTAGAGTCA	CCCTCAAAAA	TTAGTCCCAC	CTTTTCATGT	TCTATTTTTT	2550
T					2551

SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US95/13328

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C07H 21/04; C12Q 1/68; G01N 33/53; A61K 37/02; C07K 14/435, 16/18

US CL : 536/23.5, 24.31; 435/6, 7.1; 530/324, 335, 350, 388.4; 514/12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.5, 24.31; 435/6, 7.1; 530/324, 335, 350, 388.4; 514/12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS; DIALOG; MEDLINE, BIOSIS, CA, WORLD PATENTS, EP PATENTS

SEARCH TERMS: DEFENSIN, CYRPTDIN, INTESTINAL OR GASTROINTESTINAL; ANTIMICROBIAL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FEBS LETTERS, VOLUME 304, NUMBER 2,3, ISSUED JUNE 1992, OUELLETTE ET AL, "PURIFICATION AND PRIMARY STRUCTURE OF MURINE CYRPTDIN-1, A PANETH CELL DEFENSIN", PAGES 146-148, SEE FIGURE 3.	1-9, 12
X --- Y	WO, A, 93/24139 (SELSTED ET AL) 09 DECEMBER 1993, SEE PAGES 8-10 AND FIGURE 1.	1-9, 11-24, 26, 26, 27 ----- 10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	* T	later document published after the international filing date or priority date and not in conflict with the application but cited to underpin the principle or theory underlying the invention
* A* document defining the general state of the art which is not considered to be of particular relevance	* X*	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
* E* earlier document published on or after the international filing date	* Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
* L* document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	* A*	document member of the same patent family
* O* document referring to an oral disclosure, use, exhibition or other means		
* P* document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

08 FEBRUARY 1996

Date of mailing of the international search report

21 FEB 1996

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

Carla Myers

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)-w

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/13328

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ----- Y	WO, A, 93/24513 (BEVINS ET AL), 09 DECEMBER 1993, SEE PAGES 6, 7, 24, AND 25.	1-9, 11-13, 15- 17, 20, 28, 29, 36 ----- 11, 14, 18, 19, 21, 2 2, 26, 27
Y	US, A, 4,304,715 (HUDSON ET AL) 08 DECEMBER 1981, COLUMNS 6, 7 AND 12	25, 26
X ----- Y	JOURNAL OF BIOLOGICAL CHEMISTRY, VOLUME 265, NUMBER 17, ISSUED 15 JUNE 1990, OUELLETTE ET AL, "A NOVEL MOUSE GENE FAMILY CODING FOR CATIONIC, CYSTEIN-RICH PEPTIDES", PAGES 9831-9837, ESPECIALLY FIGURES 1 AND 2.	1, 12, 128, 29, 36, 37 ----- 30-35, 38

Form PCT/ISA/210 (continuation of second sheet)(July 1992)*